

AD-A263 504



AGARD-CP-525

1

AGARD-CP-525

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

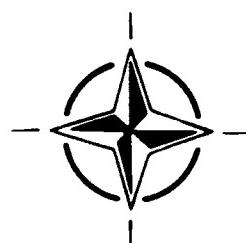
AGARD CONFERENCE PROCEEDINGS 525

Integrated and Multi-Function Navigation

(Les Systèmes de Navigation Intégrés Multifunctions)

0

*Papers presented at the Guidance and Control Panel Specialists' Meeting held
in Ottawa, Canada from 14th—15th May 1992.*



NORTH ATLANTIC TREATY ORGANIZATION

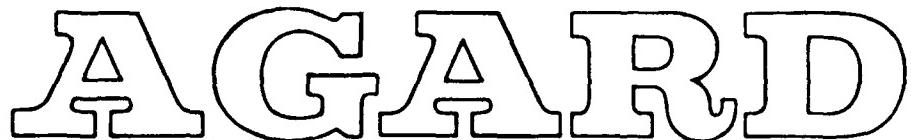


93-08674
171108

Published November 1992

Distribution and Availability on Back Cover

40th
Anniversary
Year



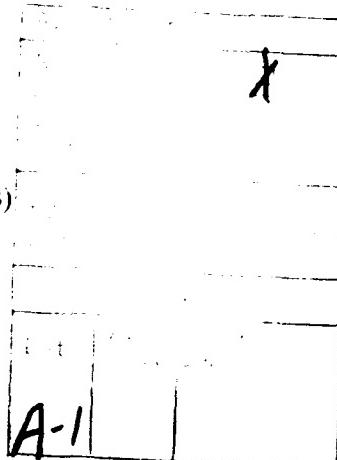
ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS 525

**Integrated and
Multi-Function Navigation**

(Les Systèmes de Navigation Intégrés Multifonctions)



Papers presented at the Guidance and Control Panel Specialists' Meeting held
in Ottawa, Canada from 14th—15th May 1992.



North Atlantic Treaty Organization
Organisation du Traité de l'Atlantique Nord

The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published November 1992

Copyright © AGARD 1992
All Rights Reserved

ISBN 92-835-0693-6



Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

Theme

Inertial navigation, blended with other navigation aids, has grown in application over the past 10 years. Blending with the Global Positioning System (GPS), in particular, has been an important development because of enhanced navigation and inertial reference performance and dissimilarity for fault tolerance and anti-jamming. Other advances include blending with radio navigation aids such as Omega and Loran-C. New concepts based upon using differential GPS and blended with inertial and visual sensors, offer the possibility of low cost, autonomous aircraft landing. Navigation using scene correlation with on-board digital maps and tracking sensors is feasible.

Multi-functional uses of inertial strapdown sensors can be achieved by addressing issues of fault tolerance (both hardware and software) and lifetime, accuracy, bandwidth, and sensor placement.

These new requirements on sensor technology have led to a number of new inertial sensor concepts emphasizing lower size, weight, power, fault tolerance, long life, and increased accuracy and bandwidth.

Thème

La navigation par inertie, intégrée avec d'autres aides à la navigation, a trouvé au cours des dix dernières années de plus en plus d'applications. L'intégration au système de positionnement global par satellite (GPS) représente, en particulier, un progrès important, puisqu'elle améliore la précision des mesures d'altitude, de vitesse et de position, la tolérance aux pannes et au brouillage. D'autres progrès ont été réalisés avec l'intégration aux aides radioélectriques OMEGA ou LORAN C. De nouveaux concepts, basés sur l'utilisation du GPS différentiel et l'intégration de mesures de capteurs inertIELS, optiques et électromagnétiques offrent la possibilité d'obtenir des systèmes d'atterrissement automatique de faible coût. La navigation par corrélation d'images grâce à des capteurs à imagerie et des cartes prévisionnelles stockées sur des calculateurs embarqués devient possible.

La réussite des applications multi-fonctions des capteurs inertIELS liés au mobile dépend des améliorations attendues dans les domaines de la tolérance aux pannes (fiabilité des logiciels et matériels), durée de vie, consommation, poids, bande passante et précision des capteurs.

Ces nouvelles exigences débouchent ainsi sur l'étude de nouvelles technologies favorisant miniaturisation, fiabilité, étendue de mesure et bande passante.

Guidance and Control Panel

Chairman: Mr S. Leek
British Aerospace
Defence Dynamics Ltd
PO Box 19
Six Hills Way, Stevenage
Herts SG1 2DA
United Kingdom

Deputy Chairman: Mr J.K. Ramage
Chief, Flight Control
Advanced Development Branch
Wright Laboratory (WL/FIGX)
Wright-Patterson AFB,
OH 45433
United States

TECHNICAL PROGRAMME COMMITTEE

Chairman:	Dr T.B. Cunningham	(US)
Members:	Dr A. Benoit	(BE)
	Dr D.F. Liang	(CA)
	ICA L. Guibert	(FR)
	Dipl.-Ing. U.K. Krogmann	(GE)
	Mr A.D. King	(UK)
	Dr J. Niemela	(US)

HOST NATION COORDINATOR

Dr David F. Liang
305 Rideau St, 7th floor
Constitution Bldg
Ottawa, Ontario K1A 0K2
Canada

PANEL EXECUTIVE

Commandant M. Mouhamad, FAF

Mail from Europe:
AGARD—OTAN
Attn: GCP Executive
7, rue Ancelle
F-92200 Neuilly-sur-Seine
France

Mail from US and Canada:
AGARD—NATO
Attn: GCP Executive
Unit 21551
APO AE 09777

Tel: 33(1)47 38 57 80
Telex: 610176 (France)
Telefax: 33 (1) 47 38 57 99

ACKNOWLEDGEMENTS/REMERCIEMENTS

The Panel wishes to express its thanks to the Canadian National Delegates to AGARD for the invitation to hold this meeting in Ottawa and for the facilities and personnel which made the meeting possible.

Le Panel tient à remercier les Délégués Nationaux du Canada près l'AGARD de leur invitation à tenir cette réunion à Ottawa et de la mise à disposition de personnel et des installations nécessaires.

Contents

	Page
Theme/Thème	iii
Guidance and Control Panel	iv
	Reference
SESSION I – MISSION APPLICATIONS	
Chairman: Dr J. Niemela (US)	
Mission Requirements and Applications of Integrated and Multi-Function Navigation Systems	1
by D.F. Liang	
Opportunities and Challenges in Avionics Integration INS/GPS – A Case Study	2
by Z.H. Lewantowicz and D.W. Keen	
SESSION II – SENSORS FOR INTEGRATED NAVIGATION AND MULTI-FUNCTION REFERENCE SYSTEMS	
Chairman: Dipl.-Ing. U.K. Krogmann (GE)	
An Overview of Optical Gyroscopes for Navigation	3
by J.G. Mark and D.A. Tazartes	
Hemispherical Resonator Gyro: Principle, Design and Performance	4
by W.W. Stripling and J.R. Baskett	
Fiber Optic Gyros and Systems	5
by H.-J. Büschelberger and E. Handrich	
Accéléromètre Silicium pour A.H.R.S. et Systèmes Hybrides de Navigation	6
par J. Leclerc, A. Defosse et O. Lefort	
SESSION III – DESIGN OF INTEGRATED AND MULTIFUNCTIONAL NAVIGATION SYSTEMS	
Chairman: Mr A.D. King (UK)	
Intégration Optimisée de l'Inertie et du GPS	7
par L. Camberlein, B. Capit et P. Debanne	
Retrofitting of GPS into Existing Navigation Suites	8
by D.I. Callender and N.F. Watson	
Integrated Precision Navigation System	9
by G. Schänzer and B. Tiemeyer	
Flight Evaluation of Differential GPS Aided Inertial Navigation Systems	10
by B.D. McNally, R.A. Paielli, R.E. Bach Jr and D.N. Warner Jr	
Visual Autonomous Automatic Landing of Airplanes	11
by E.D. Dickmanns and F.-R. Schell	
Scene Correlation for INS Aiding in Flight-Test Systems – Runway-Referenced Flight-Tests with On-Board Sensors Only	12
by B. Stieler and H.-U. Döhler	

	Reference
Cost Effective Multi-Function Integrated Avionic System by M. Silver	13†
An Hierarchic Alliance of Filters for Fault Tolerant Navigation Using Two Inertial Systems with Aiding Sensors by J.C. McMillan, J.S. Bird and D.A.G. Arden	14
Toward Achieving Global Sole Means Radionavigation Systems by F. van Graas	15
Paper 16 withdrawn	
Fault-Tolerant Multi-Function Inertial Reference Systems by M.K. Jeerage	17†

[†] Not available at time of printing.

MISSION REQUIREMENTS AND APPLICATIONS OF INTEGRATED AND MULTI-FUNCTION NAVIGATION SYSTEMS

Dr. David Feseng Liang
Directorate Research and Development Communications and Space
National Defence Headquarters
Ottawa, Ontario,
Canada K1A 0K2

1. SUMMARY

Advanced guidance, control and navigation systems are becoming more functional and integrated. This paper describes Canadian experience in the development and applications of Integrated and Multi-function Navigation Systems. The emphasis is on the development of two distinct mission-specific inertially based integrated navigation systems. In particular, in a high risk development project such as the Synthetic Aperture Radar Motion Compensation System, the special attention given to the detailed assessment of key error sources and a well-planned sequence of simulation, development and flight test evaluation are essential to the success of the project. For Multifunctional inertially-based systems, the hardware design emphasis should be on IMU architectures that minimize both the acquisition and life-cycle cost, while the software design emphasis should be on practical and efficient FDIR schemes. Some comments were also offered on the trends of GPS related future applications.

2. INTRODUCTION

Guidance, control, avionics and navigation systems for advanced military platforms and aerospace vehicles are becoming increasingly sophisticated as the demands for better mission performance and the scope of functional applications continue to escalate. Current advanced aircraft and space vehicles require multiplicity of gyros and accelerometer for automatic approach and landing; terrain following; transfer alignment; navigation; motion compensation; antenna and optical subsystem stabilization; weapon guidance and stability control.

The rapid advances of modern estimation techniques, high speed processors and the Ring Laser Gyro (RLG) technologies as well as the necessity to satisfy mission requirements in a cost/weight/reliability/performance conscious environment have led to increasing applications of integrated multi-functional, multi-sensor systems.

In view of the importance of navigation systems in contributing to successful military operations, the Department of National Defence has extensively supported the development and applications of multi-sensor integrated navigation systems as shown in Table 1. This paper will highlight Canadian experience in the development of two airborne mission-specific inertially based integrated navigation systems: the Helicopter Integration Navigation System (HINS) and the Synthetic Aperture Radar Motion

Compensation System (SARMCS). The HINS is aimed at Anti-Submarine Warfare (ASW), Anti-ship Missile Defence (ASMD) and Anti-Surface Surveillance and Targeting (ASST) missions. The SARMCS is aimed at the motion compensation of radar returns to achieve high resolution, high contrast and low geometric distortion synthetic aperture radar imagery.

Even though these two projects apply essentially the same technology and utilize similar sensors, the mission requirements are drastically different. In the case of the HINS, there were available various types of off-the-shelf navigation subsystems that could be configured to meet the specified mission requirements. However, in the case of the SARMCS, there was initially significant doubt that the mission objectives could ever be satisfied with the state-of-the-art technology, given the extremely stringent accuracy requirements specified.

This is followed by general comments on Inertially Based Multifunctional Systems and on the trends of GPS related future applications.

3. EARLY CANADIAN EXPERIENCE

In an early Canadian operational experience of an integrated airborne navigation system consisting of two inertial navigation systems, one doppler navigation system and one Omega radio receiver, the mission requirement was only specified at:

50% CEP	Radial position of 2 nautical miles with external aids of Omega.
50% CEP	Radial position error growth of 1.5 nm/hr without external aids.

The integrated system as a result of extensive flight test fully satisfied all the mission performance specifications and yet was found to have deficiency in coping with Omega lane jumps in both its Inertial/Omega/Doppler (I/O/D) and Inertial/Omega (I/O) operational modes.

As a result, it took considerable effort on the part of this author and extended flight trials to remedy the technical deficiencies in both the I/O/D and I/O operational modes. Figures 1-4 [1] illustrate the system performance before and after the software modifications, where INS 1 shows the performance of the original integrated navigation system, and INS 2 shows the performance of the modified

integrated navigation solution subject to exactly the same Omega lane jumps.

Fortunately, all our "eggs" were not in the same basket. The integrated system design was only applied to one of the two inertial navigation systems. A free-inertial solution is always available to verify the integrated solution to detect the possible influences of Omega lane jumps, and it can still be dependent to provide uncorrupted navigation data. Therefore, in both the subsequent HINS and SARMCS development projects, we have taken a more conservative approach in dedicating considerable resources to systematically simulate and study selected candidate system configurations.

4. MISSION REQUIREMENTS

The mission requirements of military aerospace vehicles vary significantly, and many of the missions must be carried out under severe dynamics and all weather and visibility conditions. For the design of any advanced integrated multi-sensor navigational system, it is important to define the detailed mission operational environment, since integrated system performance to a great extent depends on the constraints, dynamics and on the accuracy of navigation sensor in each stage of the mission. The simulated and flight test performance accuracy can only be properly measured with prescribed mission profiles. As well, the adoption of NATO STANAG 4278 95% performance accuracy measurement is essential in ensuring system performance integrity.

4.1 Missions Profiles

For the development of the HINS, two representative mission profiles for the maritime helicopter were developed for use during the simulation studies. The first profile is typical of an ASW sonobuoy mission. The HINS flight test program would seek to follow flight profiles similar to these, so that dynamic test results could be appropriately compared to simulation results.

4.1.1 ASW Sonobuoy Mission (see Figures 5 and 6).

(1) Mission Alert: (10 minutes)

- aircraft power turned on, helicopter-destroyer at speed of 18 knots. A contact is detected by the ship sensors at a range of 90 nm and a bearing at 90 degrees to the ship's course.

(2) Launch and Climb (10 minutes)

- climb at a rate of 800-1000 ft/min to a cruise altitude of 5000 ft.
- airspeed during climb 90-100 knots.
- course 235 degrees

(3) Enroute/Cruise: (45 minutes)

- course 235 degrees
- speed 140 knots, alt 5000 ft., distance flown 100-110 nm.

(4) Sonobuoy Pattern Drop: (25 min)

- drop altitude 5000 ft, speed 90 kts.
- rate 1/2 turns (1.5 deg/sec), 15 deg bank angle.
- eight sonobuoys, 5 nm spacing.

(5) Orbit/Sonobuoy Monitor: (40 min)

- orbit altitude at 5000 ft., speed 70 kts, circular or racetrack pattern, may alternate direction.
- 10 nm long orbits, 15 minute circuits.

(6) Localize/Additional Sonobuoy Drop: (15 min)

- contact passes between sonobuoys A and B.
- helicopter descends to 500 ft at 100 ft/min, flies 7 nm downrange and inserts two additional sonobuoys at C and D.
- orbit at 500 ft for 10 minutes to monitor sonobuoys A, B, C and D.

(7) Magnetic Anomaly Detector (MAD) Run: (10 minutes)

- use MAD to obtain a precise fix on the target by flying 1000 ft diameter ovals at 200 ft alt, always flying in the same direction.
- speed 90 kts, rate 1 turns (3 deg/sec), 22 degrees of bank.

(8) Attack: (5 minutes)

- helicopter descends to 150 ft and release torpedo.

(9) Transit Back to Ship: (50 minutes)

- climb to 5000 ft.
- transit speed 140 kts.

4.1.2 Convoy Screening Mission (See Figure 7)

The maritime helicopter provides close-in ASW screening support to task force operation in open ocean. Dipping sonar is used to search a moving sector formed by a 60 deg arc some 9-12 miles in front of the ships.

(1) Mission Alert: (10 minutes)

- aircraft power on.

(2) Launch/Climb/Enroute: (7 minutes)

- climb to 150 ft, transit 10 nm to the search sector and begin first sonar dip at point A.

- transit speed of 140 knots.

(3A) Sonar Dipping: (15 minute cycle)

- this dipping sequence is repeated continuously for the duration of the 4 hr mission, with sonar dip pts. at A, B and C within the sector.

- cruise velocity between dips is 90 kts.
 - cruise altitude between dips is 150 ft.
- (3B) Dipping Operation
- 700 to 1000 ft before dip point helicopter turns into wind with rate 1/2 turn, 15 deg bank, slows to 70 kts and descends to 150 ft altitude.
 - transition to the hover at 50 ft altitude (max power setting, max vibration); lower sonar, hover at 50 ft for approximately 6 minutes.
 - retract sonar, move to next dip point at 70 kts and 150 ft altitude.
- (4) Repeat Sonar Dipping Cycle 3B(b) 3B(c) above
- (5) Return to Ship: (10 minutes)
- velocity 140 kts.

4.2 Performance Accuracy Requirements

For the ASW mission the helicopter navigation system must maintain stable and accurate tactical plots over long periods of time. In the anti-surface ship targeting role, high orders of absolute and relative navigational accuracy are vital to rapid and successful action. There are further complicating factors as well. Operations must often take place under radio silence and shore-based or satellite navigation aids may be destroyed or jammed during wartime. The small crew of the helicopter must not be burdened with monitoring the functioning of, or updating, the navigation system. On the basis of full scale simulation and covariance analysis the operational accuracy specifications are established as follows:

With GPS

95% Radial position error 0.016nm (96 ft)
 95% Radial velocity error 0.47 knots
 RMS True heading error 0.1 degree at 60° Lat

Without GPS (In-Motion Alignment)

95% Radial position error 1nm/hr
 95% Radial velocity error 2.4 knots
 RMS True heading error 0.1 degree at 60° Lat

5. HINS DEVELOPMENT EXPERIENCE

For the system design of a multi-sensor maritime helicopter integrated navigation system, a large number of equipment configurations are possible. The typical approach is to use previous experience in selecting a candidate configuration in an ad hoc manner. This has the potential danger of eliminating good alternatives early in the project and could result in a suboptimal configuration. It is therefore more prudent to dedicate some resources to simulate and study two or three potential configurations with the

aim of identifying, developing and testing an integrated navigation system that best satisfies the mission requirements. The prototype of the selected system configuration can then be developed and tested with the aim of subsequent engineering development for use on board the chosen mission platform. The development effort can be divided into at least the following two phases:

5.1 Phase I - System Definition and Design.

To properly analyze the performance of various candidate configurations, a versatile simulation package was developed. Since the fidelity and performance prediction of these configurations are of prime importance, a substantial portion of the development effort was dedicated to generating complete and accurate error models. These activities are described as follows:

- a. Survey existing or soon to be available navigation sensors to determine their suitability.
- b. Identify several cost-effective candidate system configurations that can potentially satisfy the performance, weight, size and reliability requirements.
- c. Generate the navigation sensor error models.
- d. Generate a set of mission profiles to be used for trajectory generation.
- e. Develop the integration algorithms to blend the sensor outputs and undertaken performance and sensitivity analysis.
- f. Develop control and display software.
- g. Develop diagnostic software to detect sensor failure.

In the detailed simulation analysis both covariance analysis and Monte Carlo simulation software were used. Covariance analysis software is effective for the design of integration filters because it provides ensemble statistical data. Because ensemble statistics are the outputs, covariance analysis can (in many circumstances) reduce the need for computationally intensive Monte Carlo simulations. This statistical information can be used to assess candidate Kalman filter designs and to project the performance of a particular navigation system configuration. Covariance analysis is very useful for assessing the effects of mismatch between the filter design model and the "real world" or truth model. In addition, it can be readily used to establish error contribution tables and error budgets which let the filter designer focus in on the major error contributors.

However, Monte Carlo simulation also has its place in the design of suboptimal Kalman filters. Simulation can be particularly effective in assessing the effects of nonlinearities which are difficult to address in the covariance analysis framework. Also,

certain types of mismatches between the Kalman filter model and the truth model are more conveniently addressed with the simulation program than with the covariance program. An example of this is the sensitivity assessment of the effect of sea current correlation time mismatch.

Another area in which the Monte Carlo simulation can be of more use than the covariance analysis program has been in the investigation of the effects of unmodelled manoeuvre-dependent sensor errors. As well, covariance analysis software will never be able to replace the function of simulation software for final checkout of Kalman filter code.

5.2 System Error Modelling

Detailed simulation and error models for all the relevant navigation equipment and environmental disturbances were developed for the performance evaluation and sensitivity analysis.

A wide variety of types and brands of navigation sensors can contribute to meeting a specific set of requirements. The following list of generic navigation subsystems can be considered having merit in HINS applications:

- Global Positioning System (GPS),
- Inertial Navigation System (INS),
- Attitude and Heading Reference System (AHRS),
- Doppler Radar,
- TACAN,
- Omega,
- Air Data System,
- Strapdown Magnetometer,
- Radar Altimeter.

However, many of the generic equipment categories can be further subdivided. For example, in the INS category there are high and medium accuracy inertial equipments available. The inertial sensors can be mechanized in either a gimballed or strapdown platform. Strapdown configurations may use conventional or ring laser gyroscopes. Similarly, Doppler velocity sensors and Doppler navigation systems are available in the Doppler category.

5.3 Phase II Development and Testing

- a. Acquire the sensor hardware for the preferred configuration and complete the development of the design produced in Phase I.
- b. Develop the data bus and interface electronics to connect the sensors.
- c. Refine and convert the preliminary forms of the Phase I Kalman filter and integration software into real-time flight implementable programs. Develop the control and display software and transfer these programs into the airborne processor.
- d. Conduct static tests of the completed system in an integration laboratory and in low dynamics tests using a mobile van.

- e. Conduct flight tests in a test aircraft to evaluate system performance in a realistic vibration environment.
- f. Assess the navigation system's performance by conducting extensive flight tests in a test aircraft on an instrumented range.

The Phase II development was aimed at realizing in hardware the Phase I system design by constructing and developing through ground and flight testing the Advanced Development Model (ADM) to produce a fully developed and flight validated integrated navigation system prototype.

5.4 HINS Configuration

The selected HINS configuration is represented in Figure 8. The navigation bus interfaces the primary sensors (INS, GPS, DVS) with the navigation computer system (NCS), which contains the Integrated Nav Function, described below. The AHRS interfaces directly to NCS via an ARINC 429 interface. The Control Computer System (CCS), which serves as a smart interface between the NCS and the Control and Display Unit (CDU), is connected to the NCS by a second 1553B data bus.

The main hardware items are:

- a. Honeywell H-423 Standard Ring Laser Gyro Navigation System.
- b. Rockwell-Collins model 3A NAVSTAR Global Positioning System Receiver.
- c. Canadian Marconi ANP-235 Doppler Velocity Sensor (DVS).
- d. HINS MIL-STD-1553B Navigation Bus.
- e. HINS Motorola 6800 Microprocessor Navigation Computer System.
- f. HINS Control and Display Computer Subsystem.
- g. Data Loader System/Data Loader Module.
- h. HINS Control Data Bus.

5.5 Flight Test Summary

The HINS ADM was flight tested at Holloman Air Force Base, New Mexico, and Vandenberg Air Force Base, California, to take advantage of an accurate, Completely Integrated Reference Instrumentation System (CIRIS). The flight tests were conducted over land and sea with prescribed mission profiles. At Holloman AFB, three types of alignments (ground, in motion, and enhanced-interrupted) four operational modes (Full-Up, INS/DVS, GPS/DVS, AND INS/GPS) and three profiles (high dynamic, low dynamic and sonobuoy) were flight tested.

At Vandenberg AFB, two different types of alignments (in-motion and ground), five operating modes (Full-up, INS/DVS, GPS/DVS, DVS and GPS) and two profiles (sonobuoy and racetrack) were flown.

5.5.1 Over Land Flight Tests

At Holloman AFB, the ensemble results [2] of 4 two and a half hour full-up mode are presented in Table 2. The 95th and 50th percentile radial velocity and operational flights position error plots are presented in Figures 9 and 10.

For the TNS/DVS mode with in-motion alignment, Table 3 presents the ensemble results of 4 flights with a average alignment time of 23 minutes and navigation mode operations of 72 minutes. The composite radial velocity and position errors are presented in Figures 11-12. Zero time for the plots was the point at which the inertial system was switched to the "navigate" mode.

5.5.2 Over Sea Flight Tests

At Vandenberg, the HINS was flown over water with in-motion alignment of 40 minutes before switching into the navigation mode for an average of 100 minutes. The ensemble results [2] of 4 flights are presented in Table 4. The composite radial velocity and position error plots are presented in Figures 13 and 14. Sharply elevated position errors are present at the end of three flights apparently due to sea-to-land transitions.

6. SYNTHETIC APERTURE RADAR MOTION COMPENSATION (SARMCS) EXPERIENCE

To attain high resolution, high contrast and low geometric distortion airborne imagery for the AN/APS-506 radar, it is essential that very accurate motion compensation be applied to the radar returns to account for deviations of the radar phase centre from a smooth reference path, chosen *a priori*. The main task of the SARMCS is to determine spurious high frequency deviations from the desired motion along the reference path; these displacements are used by the SAR processor to adjust the relative alignment and phase of the radar returns, after which it removes low frequency (quadratic) errors by autofocusing, which amounts to fitting a quadratic phase adjustment to the radar returns across the synthetic aperture to maximize a specific measure of image contrast.

6.1 Motion Compensation Requirements

There are two main modes of SAR operation. In spotlight mode, the antenna is aimed at the designated target (Figure 15) or specified coordinates, and wide bandwidth radar pulses are emitted. In the strip-mapping mode, the orientation of the radar boresight is held constant approximately at right angles (Figure 16) to the nominal flight path, thus illuminating a swath to the side of the aircraft.

For airborne SAR processing, the ideal situation is that the radar antenna, mounted on the aircraft, moves along a straight line in space, transmitting and receiving pulses at equally spaced intervals along this path, which forms the synthetic aperture. However, in general, the actual path of the antenna will

deviate from the nominal path due to aircraft turbulence, autopilot inaccuracies, etc. These spurious motions, if uncompensated, can severely degrade the SAR image. The function of the motion compensation system is to sense the antenna motion, and compute the deviations between the actual path and the nominal path. This information is then used to correct the phase of the radar returns so that, ultimately, as far as the radar processor is concerned, the pulses look like they were emitted at these ideal points along the nominal track.

6.2 Performance Accuracy Requirements

Since this project is in support of the development of a SAR radar capability, the performance requirement for motion compensation has been specified in terms of power spectral density (PSD) of the tolerable error in measurement of the displacement of the antenna phase centre along the radar line-of-sight (LOS) (Figure 17). The displacement error spectrum has been divided into two components. The portion of the PSD below A Hz lies in the "Don't-Care Region". This contains the components of the displacement error which have characteristics times longer than the maximum aperture time of T seconds ($A\text{ Hz} = 1/T$ seconds). Displacement error components having frequencies above A Hz must be controlled by the motion compensation system. An estimate of the RMS magnitude of the allowable displacement error measured over the scene is:

$$\delta r = 0.33 \text{ mm (RMS)}.$$

This is the residual displacement error along the LOS to the target not including contributions from constant and linear components of displacement error, which have no effect on image quality, or quadratic components which are removed by the autofocusing.

The displacement error components above A Hz affect the contrast of the SAR image while those below A Hz degrade resolution and produce geometric distortions of the image. Although the lower frequency components within the "Don't-Care" region may affect the SAR image quality, this error component is to be controlled through the autofocusing algorithm. Therefore the SARMCS project only considers error sources above the A Hz region. It should be emphasized that during the early phase of the design study, it was felt that the performance requirement of the order of a millimetre would not likely be achieved in a practical situation as there are a large number of practical design variables which could easily overwhelm this level of accuracy requirement.

6.3 Analysis of Error Sources

In view of the extremely stringent accuracy requirement, primary design consideration was concentrated on estimating the magnitude of the contribution of major system error sources to the residual LOS displacement error. The following error sources are considered:

- Uncompensated phase centre motion due to antenna pointing errors

- computational and related errors arising from the imperfect solution of the strapdown motion compensation equations
- gyro and accelerometer errors of the strapdown IMU
- attitude errors of the strapdown navigator.

The analysis indicated that the principal sources of error in computing the LOS displacement are the mislevels of the strapdown navigator, the error in determining the initial depression angle of the target LOS, computational errors in navigation and targeting algorithms, error in the measured relative azimuth of the strapdown IMU and the radar antenna boresight, and errors in the accelerometers and gyros. From the results obtained for the required aperture under conditions of medium turbulence, the error budget of Table 5 was drawn up.

6.4 SARMCS System Configuration

The SARMCS instrumentation that was designed and installed on board the National Aeronautical Establishment (NAE) Convair 580 aircraft for Phase II is shown in Figure 18. The system configuration consists of:

- a. a ring laser gyro inertial navigation system, denoted the master INS, which is located about three meters from the aircraft centre of gravity.
- b. a Doppler radar employing a strapdown three-beam lambda configuration.
- c. a specially designed strapdown inertial measurement unit, denoted as MCIMS (Motion Compensation Inertial Measurement Subsystem), which is located about 50 centimetres from the SAR antenna phase centre.
- d. air data sensors including a static air pressure transducer along with an air temperature probe to determine barometric altitude.

The simplified block diagram of Figure 19 shows the SARMCS functions. All of the blocks within the dotted line are software modules which together comprise the motion compensation processor. The primary sensor is the MCIMS which accurately measures rotational and translational motions. This unit is mounted on the SAR antenna in the nose of the aircraft to provide as direct a measurement as possible of the antenna motion. The raw measurements from the strapdown unit are processed in a strapdown navigator algorithm to yield antenna position, velocity and attitude. This information is then used in a targeting algorithm to generate motion corrections for the radar returns. These include, in addition to phase corrections, adjustments to the radar pulse repetition frequency and range gate slewing to account for aircraft motion. The air data is used to compute baroaltitude which is needed to stabilize the vertical channel in the strapdown

navigator. Outputs from a Doppler velocity sensor and a master inertial navigation system on board the aircraft are fed into a Kalman filter algorithm along with outputs from the strapdown navigator. The Kalman filter optimally integrates this information and estimates the errors in the various sensors. The net effect is a transfer of alignment from a Doppler-damped master INS to the strapdown navigator. The error estimates for the strapdown navigator are fed back into the strapdown algorithm and used there to correct the relevant parameters. This error control scheme utilizing the Kalman filter is designed to prevent a build-up of long-term levelling errors commonly referred to as tilts in the strapdown platform which, as indicated in Table 5, are important contributors to motion compensation errors.

6.5 SARMCS Simulation Software

The SARMCS simulation software is divided into 3 separate packages:

- a. The Data Synthesis Package generates realistic synthetic data from the system error models which include:
 - strapdown IMU model
 - Master INS model
 - Doppler radar model
 - Atmospheric pressure and temperature models
- b. The Data Processing Package can process both synthetic and real sensor data. It implements the motion compensation processor of Figure 19, which includes:
 - strapdown navigation and sensor compensation algorithms
 - barometric altitude algorithm
 - Kalman filter
 - targeting algorithms
- c. The Evaluation Package evaluates the performance of the processing package by comparing computed master and strapdown navigator positions, velocities and attitudes with corresponding accurate reference data generated by the Synthesis Package.

6.6 Kalman Filter Design

A baseline 35 state Kalman filter was developed to indicate the best level of achievable performance. The filter is mechanized using Bierman's U-D factorized formulation [3]. The filter structure is implemented in such a way that arbitrary subsets of the full error state vector may be selected for a particular run by specifying values for input parameter tables. Similarly, any subset of measurements may be selected. This results in a very flexible design tool.

An error control routine uses position, velocity and misalignment error state values to correct position, velocity and attitude estimates of the strapdown navigator after each filter update, unless the update occurs during a SAR window. The corresponding error states are zeroed after executing error control.

For practical implementation, a suboptimal 21 state Kalman filter was designed. The error state vector is described in Table 6.

The modelling of relative strapdown (S/D) system errors instead of absolute S/D system errors in the state vector is a design decision that is motivated by several considerations. First, from a theoretical viewpoint, this is an appropriate choice because measurements constructed by comparing information from two systems with the same error dynamics only allow observability of the relative error between the two systems. A practical motivation for modelling relative S/D errors is that for this case where S/D instrument errors are expected to be much larger than master INS instrument errors, it can be shown that the estimation of x_v and x_w is essentially decoupled from the estimation of x_m , x_{m_d} , x_d , in the sense that no significant correlation develops between these two sets of subvectors during Kalman filter operation. This behaviour is mathematically equivalent to having two independent Kalman filters, one of which accomplishes transfer-of-alignment from the master to the S/D platform while the other performs Doppler-damping of master errors. This is a robust configuration in that the effects of slight mismodelling of the lower quality S/D IMU in the Kalman filter cannot feed back through the velocity matching measurements to corrupt the estimation of master system errors.

One feature of the SARMCS Kalman filter that noticeably distinguishes it from a navigation-type Kalman filter is the absence of measurements that bound the inertial position error. This is a direct consequence of tailoring the Kalman filter for the specific task of performing accurate S/D platform alignment. The presence of S/D position errors has only a relatively weak effect on the buildup of platform misalignments and velocity errors, so there is no need to accurately estimate them in the SARMCS. In fact, S/D position states are dropped altogether; the use of the ϕ angle error formulation for the Kalman filter inertial error models conveniently allows this to be done without impacting on the filter's ability to estimate S/D velocity errors and platform misalignments from the velocity matching measurements.

It is worthy to note that the choice of augmenting states in the SARMCS Kalman filter is based primarily on the criterion of observability. The states in x_v , x_d and x_w represent only those significant instrument errors that are separately observable with the given measurements.

6.7 SARMAC Simulation and Flight Test Verification

The mission profile for the simulation is shown in Figure 20. It is consistent with the flight tolerance limits of the CV 580 research aircraft. It involves an initial climb to an altitude of 1000 metres, followed by a racetrack manoeuvre after 10 minutes and an s-turn 20 minutes after takeoff. This is followed by a period of nominally straight and level flight during which 23 SAR apertures are simulated. About 1 hour after takeoff, a second s-turn is carried out to control the strapdown navigator heading error. This is followed by another section of straight and level flight during which 9 more SAR apertures are simulated. All simulation results are obtained using the baseline Kalman filter design unless otherwise stated.

6.7.1 Simulation Results

Figures 21 to 23 show the north velocity, roll and heading errors of the strapdown navigator, together with RMS values computed from the Kalman filter error covariance. Notice that the σ bound of the roll error conforms well with the single run error trace. In Figure 23, the heading error RMS is reduced to the level of master heading error after the second s-turn.

Simulation results indicating the performance of the suboptimal 21 state filter are shown in Figure 24. This plot depicts the roll errors of the strapdown navigator, as well as the filter predicted 1σ value. From the comparison of Figure 24 to Figure 22, the performance of the two filters is quite similar except for brief periods during aircraft manoeuvres.

6.7.2 Flight Testing Philosophy

The philosophy for flight testing the SARMCS system involves validating the correct operation of subsystem configurations which increase in complexity until the complete configuration is attained. This type of approach provides a systematic method for detecting and isolating unexpected error sources in the hardware and/or software functions. There are five sequential steps in this flight test plan:

- 1) evaluate the performance of the Master/baro subsystem,
- 2) evaluate the performance of the MCIMS/baro subsystem,
- 3) evaluate the performance of the Master/Doppler/baro subsystem,
- 4) evaluate the performance of the Master/Doppler/baro/MCIMS subsystem,
- 5) evaluate the performance of the full SARMCS with all SARMCS sensors utilized.

In the first four steps, the subsystems are tested to verify that their velocity accuracies are consistent with expected values predicted by

the earlier simulations. A flight reference system (FRS) is used to provide the "truth" data for evaluation of these subsystems. The FRS employs an extended Kalman filter which optimally integrates precision microwave ranges from a Del Norte Triponder system with information from an LTN-91 inertial navigation system. For the FRS, three ground transponders are positioned at surveyed locations. They provide a rectangular coverage area of width 45 kilometres and length 130 kilometres within which at least two ranges with good geometry are received by the aircraft. Under these conditions, the FRS provides continuous aircraft velocity information accurate to 0.1 metres/second and position information accurate to 10 metres. This accuracy is sufficient for evaluating these subsystems since velocity errors in the order of 1 metre/second are expected.

After it has been verified that the various subconfigurations are operating properly, the final step involves testing the full SARMCS configuration by applying motion corrections computed by the SARMCS system to spotlight SAR data. The extent to which the SAR image is enhanced is the ultimate indication of the performance of the SARMCS.

6.7.3 Flight Test Results

Currently, the motion compensation system is already implemented in real-time. However, the radar data are recorded on high density digital tape, for processing in a ground-based facility. As a first step in the processing, motion compensation vectors are generated and applied to the radar data, and then the actual SAR processing is done.

Figures 25 to 28 indicate the performance of the SARMCS using the suboptimal Kalman filter. Figure 25 shows the motion compensated SAR image of essentially a point scatterer. The target is a satellite receiving station's 10 meter reflector-type antenna, equipped with a dual frequency S/X-band feed. The X-band feed was short circuited in order to provide a strong reflection from the antenna. In this type of display, the vertical scale is signal amplitude, one horizontal scale is distance along the radar line-of-sight, and the other horizontal scale is distance perpendicular to the radar line-of-sight. Theoretically, the SAR image for a point target should be one sharp peak, which is fairly close to what is being achieved in this image.

Figure 26 shows the same radar data but processed without motion compensation. Here, it is apparent that the effect of spurious uncompensated aircraft motion is to cause energy from the main peak to spill into the side lobes. The implication of this for a more typical image containing many point targets is that weaker targets in the vicinity of stronger ones might be completely obscured by this sidelobe energy; in optical terms, the resulting image would be described as having poor contrast.

Figure 27 shows a motion-compensated SAR strip map image taken of the Sudbury area in northern Ontario. The strip is about 650 metres

wide and 1500 metres long. The area is near the nickel smelter in Sudbury that refines the ore from the surrounding mines. The dim lines are roads, and the brighter ones are railway tracks. The very bright lines are pipes that lead into a cluster of five circular storage tanks; the radar reflection off the circular edge of one of the tanks can be clearly seen in the image.

Again, Figure 28 shows the same piece of radar data but processed without motion corrections. The image is initially focused but then starts to smear out as the aircraft deviates from the nominal straight line track. In this case, the deviation is a result of a 2 degree change in aircraft heading. The pipelines can barely be distinguished in this image.

7. INERTIALLY-BASED MULTIFUNCTIONAL SYSTEMS

Historically, the body mounted flight control gyros and accelerometers were not of sufficient accuracy for navigation. The navigation sensors, on the other hand, were more accurate than necessary and yet not suitable for flight control since they were gimballed and did not provide the platform angular rate and acceleration. Recent advances in Ring Laser Gyro and high speed microprocessors have made advanced inertially-based multifunctional systems feasible and desirable. The RLG has both the accuracy required for navigation and the dynamic bandwidth required for flight control. Microprocessors have enabled the processing of fault detection, isolation, dynamic reconfiguration, flight control compensation and redundancy data management functions to achieve fault-tolerant multi-functional operations.

7.1 Performance Accuracy

The navigation performance requirements for Multifunctional Inertial Systems are in general, similar to medium accuracy inertial system. Table 7 represents typical system output accuracies that are required. The anticipated one sigma sensor error budgets are shown in Table 8.

Inertial navigation gyros and accelerometers are normally orders of magnitude more accurate than those commonly used for flight control. Flight control sensor accuracy requirements (Table 9) are much less stringent. The paramount requirements are mount characteristics, sample rates, structural bending, signal to noise ratio, and aliasing effects and computational delays. Therefore, no software compensation on the sensor errors will be required, due to the superior quality of inertial grade sensors.

7.2 Redundancy Requirements

Fault tolerant architectures with redundant inertial sensors are increasingly required to meet the operational requirements of high reliability launch vehicles, advanced aircraft, maintenance free commercial aircraft and long-life satellites, etc. Traditionally, fault tolerance is provided through hardware

redundancy. Triple and penta-triad architectures have been adopted for commercial aircraft and space vehicles. Such architectures lead to a large number of sensors and associated cost, weight, size and power considerations. However, high redundancy and reliability can also be achieved through a minimum set of redundant components in an integrated multifunctional design with fault detection and isolation (FDI) management functions. In general, an integrated flight control system must ensure system integrity to provide fail-operate, fail-operate and fail-safe (FO, FO, FS) operations. This means that the aircraft must possess a full authority flight control system after two component failures. Even after three component failures, it still must operate safely but with degraded handling qualities.

7.3 Survivability

Flight control sensors are not considered survivable when the gyros are clustered at a single location near the primary aircraft bending antinode or when the accelerometers are clustered at a node. To improve survivability of aerospace vehicles, it is necessary to decrease the vulnerability of critical flight control sensors to penetrator or explosive warheads. This requires assurances of a) no single point vulnerability, b) the flight control output will survive single-hit battle damage from one 23mm High Explosive Incendiary (HEI).

7.4 Mission Reliability

Reliability must be considered in every phase of the design process. High mission reliability can be achieved either by designing components that have a very large MTBF (mean time between failure) or by providing redundant components and a redundancy management scheme so that performance goals are met in spite of the failures.

System reliability measures the degree to which fault tolerance is being achieved. It can be used as a yard stick to compare different fault tolerant architectures. Reliability analysis of a number of fault tolerant IMU architectures with redundant inertial sensors was performed by Jeerage [4]. He adopted a systematic reliability analysis procedure to compare this state-of-the-art fault tolerant IMU architectures. However, if any inertially-based multifunctional systems were to be adopted, it would be necessary to demonstrate on an economic term the life cycle cost benefits of fault tolerant architectures. The mission reliability data alone will never be sufficient to demonstrate its operational applicability.

7.5 FDIR Management

The design of any failure detection isolation and reconfiguration system should consider the following design implications:

- a) **Failure detection:** Rapid response in detecting failed or degraded sensors. This requires the system to have a large band-

width so as to be sensitive to high frequency effects. However, higher frequency components can cause undesirable flight control aliasing effects such as flutter and pilot discomfort. As well, increasing the bandwidth of the Kalman filter, will make the overall system more sensitive to noise, thus degrading the performance under no-failure conditions. As well a false alarm can cause the removal of a good sensor leading to erroneous detection of follow-on sensor failures. Therefore the FDIR must be efficient and real time implementable, while addressing these conflicting requirements; and

- b) **Failure Isolation:** Upon the detection of a failure, isolation is required to identify the failed sensor. With a wrong isolation, a good sensor can be eliminated, while leaving the failed sensor to provide the system performance and FDI measurements. Aside from direct built-in test and verification of reasonableness, and consistency, numerous papers have been published on FDI algorithms, they generally fall into the following categories:
 - i) **Voting Scheme:** This is relatively easy to implement and provide fast detection and isolation of hard failures. It is often useful in systems of high hardware redundancy. A majority voting technique is generally used to detect failures and eliminate faulty instruments. However, it cannot take advantage of redundant sensor information and requires multiple hardware redundancy. Kerr [5] has proposed a voting scheme based on the pair wise statistical comparison of the solutions from a series of sensors. Any lack of conformity between two solutions are sent to a voting scheme for possible isolation of a sensor failure.
 - ii) **Generalized Likelihood Test:** This approach employs a set of equations called the parity equations to detect a failed sensor. A parity equation is defined as the comparison of the measurements between several sensors such that the result is zero, except for a small noise residual. A significant increase in the residual measurement indicates a failure due to a shift in sensor bias. To take full advantage of this technique, sensor geometry must be configured to yield failure detectability. This is the most widely used FDI techniques, since fault detection and isolation tasks are separately performed, and it has been shown to be equivalent to other methods such as the total squared error (TSE)/maximum square residual (MSR) approach [6].
 - iii) **Multiple-Decentralized Filters:** Recently, a number of multiple decentralized filtering schemes have been introduced in literature.

Bierman [7] has developed a flexible decentralized filter. Brumback and Srinath [8] proposed the computation of multiple (local) estimates conditioned on different subsets of the sensors. A statistical overlap test is used to determine the validity of the local estimates, and a failed sensor is identified from analysis of the invalid local estimates.

Carlson [9] has developed a federated filter by extending the work of Bierman providing usable and physically meaningful local filters and allows master filter rate reduction via local data compression (prefiltering). The accuracy is dependent on the selectable master filter rate, and provides a high degree of fault tolerance. In Canada, we have attempted to investigate three FDIR techniques to detect soft failures in multi-sensor integrated navigation system. These are:

- a) Residual test,
- b) Chi-square state-vector test, and
- c) multiple model filtering based on a Bayesian fault isolation scheme.

The results of this study will be presented by McMillan, Bird and Arden in this Specialist's Meeting. The success of any FDIR schemes must be measured in terms of:

- a) Computational complexity,
- b) False alarm minimization,
- c) Failure detection sensitivity, and
- d) Optimal performance accuracy and graceful degradation.

8. GPS RELATED FUTURE APPLICATIONS

The success of Desert Storm operations, amongst other things, convincingly demonstrated the utility of GPS and scene correlation/terrain reference systems and smart munitions.

Recent advances in mobile communication technology will usher in abundant opportunities to integrate communications, navigations and position reporting systems. Highly accurate differential GPS will find greater applications in constricted waterways, terminal guidance, precision approach and stand-off weapons. The integration of GPS with scene correlator will no doubt enhance accuracy of terminal guidance systems, mission planning efficiency and terrain following/avoidance applications. This type of integrated system should also find significant applications in low flying aircraft and aerial surveillance systems.

9. CONCLUSIONS

The advent of strapdown RLG, GPS, terrain reference systems, multi-sensor integration filter and fault tolerant FDIR techniques has extended applications of IMU and GPS far beyond the traditional navigation domain. The successful development and applications of these technologies in military operations will no

doubt usher in abundant opportunities for civilian aerospace, marine and land applications and for wide ranging innovative military applications. Even for the system design of highly accurate performance requirement, the design risk can be adequately managed with a well executed sequence of simulation, development and flight test evaluation activities.

REFERENCES

- [1] Maritime Air Group Final Report on OPVAL/A1033, 26 September 1991.
- [2] Menefee, L. and Hunt, C. "Evaluation Flight Testing of Helicopter Integrated Navigation System", AFDTTC-TR-91-57, October 1991.
- [3] Bierman, G.J., Factorization Methods for Discrete Sequential Estimation, New York: Academic Press, 1977.
- [4] Jeerage, M.K., "Reliability Analysis of Fault-Tolerant IMU Architectures with Redundant Inertial Sensors", 1990 IEEE Position Location and Navigation Symposium, pp 587-592, 1990.
- [5] Kerr, T., "Decentralized Filtering and Redundancy Management for Multisensor Navigation", IEEE Transactions on Aerospace and Electronic Systems, vol. 23, no. 1, January 1987, pp. 83-119.
- [6] Daly, K.C., etc., "Generalized Likelihood Test for FDI in Redundant Sensor Configurations", AIAA Journal of Guidance, Control and Dynamics, vol. 2, no. 1, Jan-Feb. 79.
- [7] Bierman, G.J., and Belzer, M., "A Decentralized Square Root Information Filter/Smoother", 24th IEEE Decision Control Conference, Fort Lauderdale, FL, December 1985.
- [8] Brumbach, B.D. and Srinath, M.D., "A Fault-Tolerant Multi-sensor Navigation System Design", IEEE Trans on Aerospace and Electronics Systems, Vol-123, no. 6, November 87, pp. 738 - 756.
- [9] Carlson, N.A., "Federated Square Root Filter for Decentralized Parallel Processes", IEEE Transactions on Aerospace and Electronic Systems, vol. 26, no. 3, Mar 90, pp. 517-525.

TABLE 1
**CANADIAN DEVELOPMENT AND APPLICATIONS OF
MULTI-SENSOR INTEGRATED NAVIGATION SYSTEMS**

<u>DR SENSORS</u>	AIR			MARINE		LAND
	CP140	HINS	SARMCS	MINS	DIINS	PLANS
Speed Log				X	X	
Air Data Sensor	X	X				
Gyro Compass				X		X
Magnetometer						X
Odometer						X
Doppler Sensor/ Navigator	X		X			
AHRS		X	X			
Inertial Navigation (INS)		X	X			
Dual INS	X				X	
<u>RADIO</u>						
Loran C				X	X	
Omega	X			X	X	
Transit				X	X	X
GPS		X	X	X	X	X

TABLE 2
**ENHANCED INTERRUPTED ALIGNMENT,
FULL UP MODE (95TH PERCENTILE)**

TIME (MIN)	POSITION ERROR (NM)	SPEC (NM)	VELOCITY ERROR (KNOTS)	SPEC (KNOTS)
60	0.011	0.016	0.26	0.47
115	0.010	0.016	0.27	0.47

TABLE 3
**IN-MOTION ALIGNMENT, INS/DVS
OVER LAND (95TH PERCENTILE)**

TIME MIN	POSITION ERROR RATE (NM/HR)	SPEC (NM/HR)	VELOCITY ERROR (KNOTS)	SPEC (KNOTS)
60	1.45	1.00	0.94	2.4

TABLE 4
**IN-MOTION ALIGNMENT, INS/DVS
OVER WATER (95TH PERCENTILE)**

TIME (MINS)	POSITION ERROR RATE (NM/HR)	SPEC (NM/HR)	VELOCITY (KNOTS)	SPEC (KNOTS)
60	0.19	1.00	2.26	2.40
90	0.67	1.00	2.28	2.40

TABLE 5
ERROR BUDGET

ERROR SOURCE	MAX CONTRIBUTION TO RMS (MM)	%
1. Sensor errors		
- accelerometer	0.02	0
- gyro	<u>0.09</u>	8
RSS TOTAL (sensors)	0.09	8
2. Computational errors	0.10	9
3. S/D heading error	0.02	0
4. S/D mislevels	0.18	30
5. S/D azimuth alignment	0.05	2
6. Azimuth angle encoder	0.05	2
7. Interpolation errors	0.10	9
8. Initial depression angle	0.15	20
9. Contingency	<u>0.15</u>	<u>20</u>
RSS TOTAL (all)	0.33	100

TABLE 6
KALMAN FILTER ERROR STATES

X, Y	MASTER POSITION ERRORS
X, Y	MASTER VELOCITY ERRORS
X, Y, Z	MASTER PLATFORM MISALIGNMENTS
X, Y	MASTER ACCELEROMETER BIASES
X	DOPPLER SCALE FACTOR ERROR
Z	DOPPLER BORESIGHT MISALIGNMENT
X, Y	SEA CURRENT BIASES
X, Y	RELATIVE S/D VELOCITY ERRORS
X, Y, Z	RELATIVE S/D PLATFORM MISALIGNMENTS
X, Y, Z	S/D GYRO BIASES

TABLE 7
NAVIGATION PERFORMANCE REQUIREMENTS

Radical Position Error Rate	1nm/hr (CEP)
Velocity Error	3 ft/sec (RMS)
Reaction Time	5 minutes

TABLE 8
NAVIGATION SENSOR REQUIREMENTS (ONE SIGMA)

Gyro:	
Bias	0.008 deg/hr
Scale Factor	5 ppm
Alignment	25 μ rad
Random Noise	0.002 deg/sqr hr
Turn-on Transient	0.007 deg/hr
Accelerometer:	
Bias	40 μ g
Scale Factor Stability	35 ppm
Alignment	40 μ rad

TABLE 9
FLIGHT CONTROL SENSOR REQUIREMENTS

Gyro:	
Bias	1.5 deg/hr
Scale Factor	0.1%
Alignment	1 millirad
Resolution	0.02 deg/sec
Range	400 deg/sec
Accelerometer:	
Bias	4 milli-g
Scale Factor	0.1 %
Alignment	7 millirad
Resolution	2 milli-g
Range	20 g

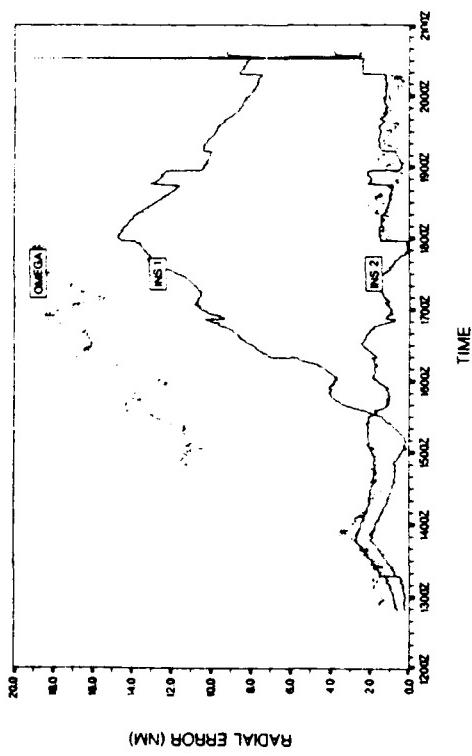


Figure 1 : Integrated INS/Omega Mode Performance

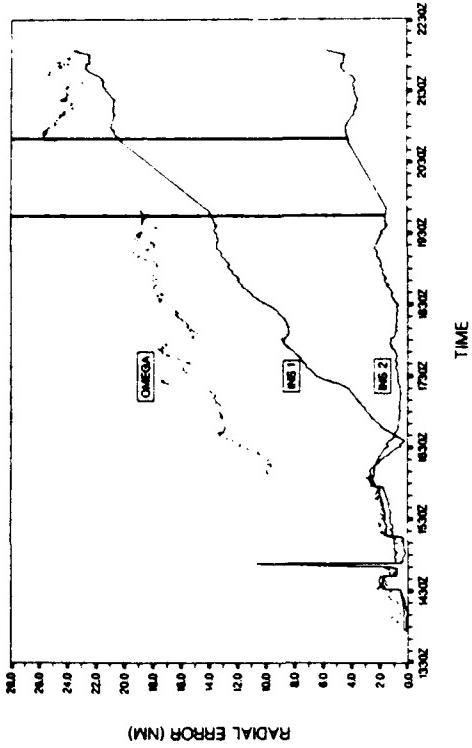


Figure 2 : Integrated INS/Omega Mode Performance

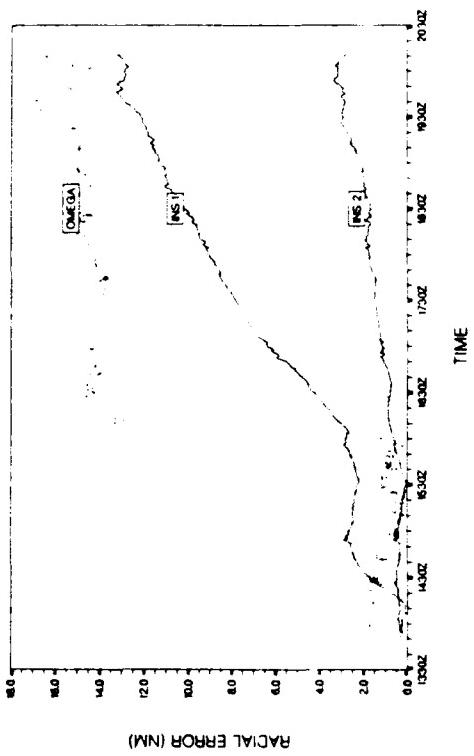


Figure 3 : Full-Up Integrated I/D/O Mode

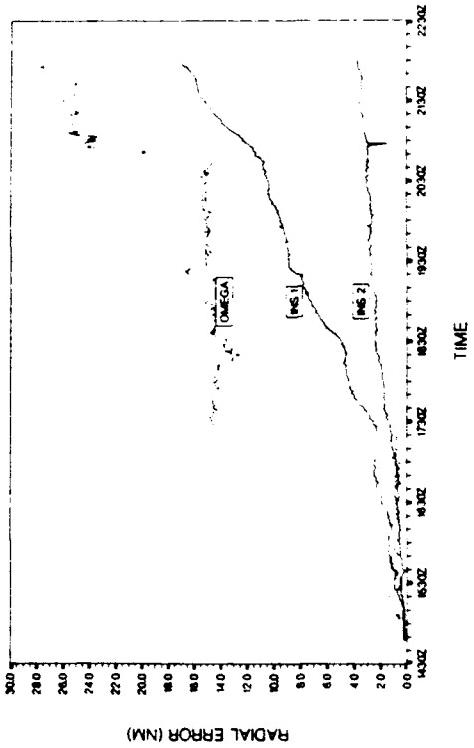


Figure 4 : Full-Up Integrated I/D/O Mode

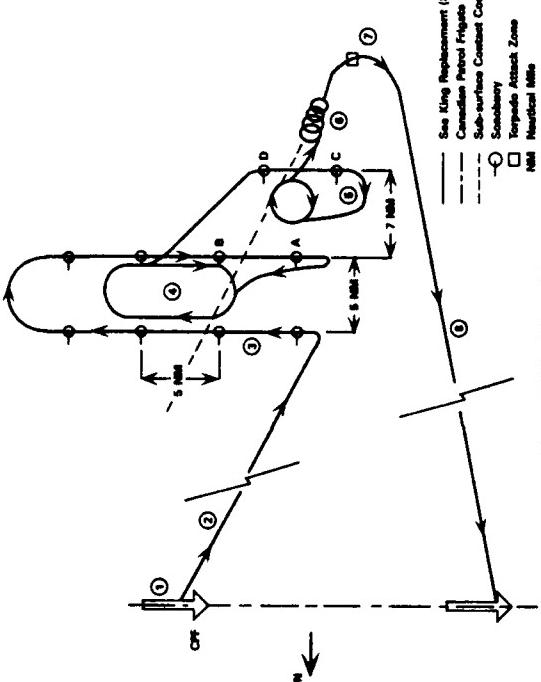


Figure 5 : Plan View of Sonobuoy Mission

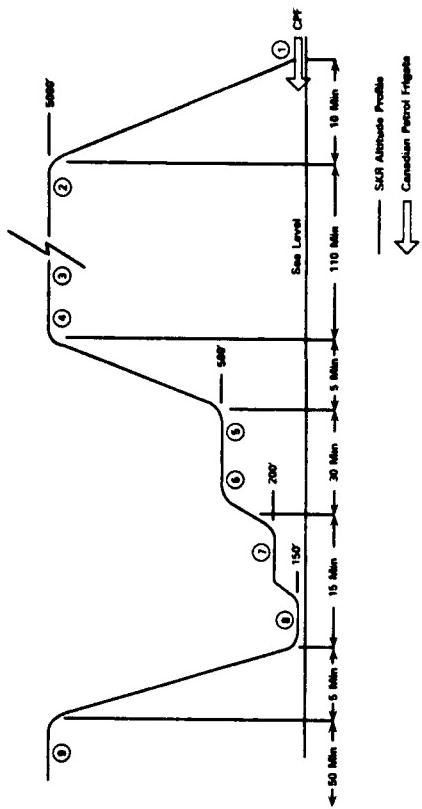


Figure 6 : Altitude Profile for Sonobuoy Mission

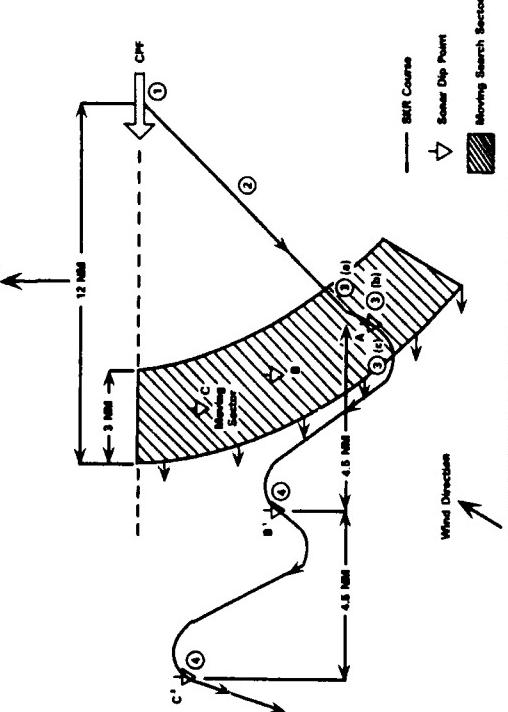


Figure 7 : Plan View of Convoy Screening Mission

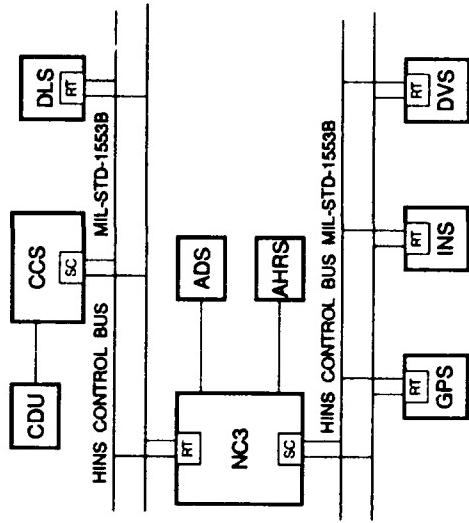


Figure 8 : HINS ADM Architecture

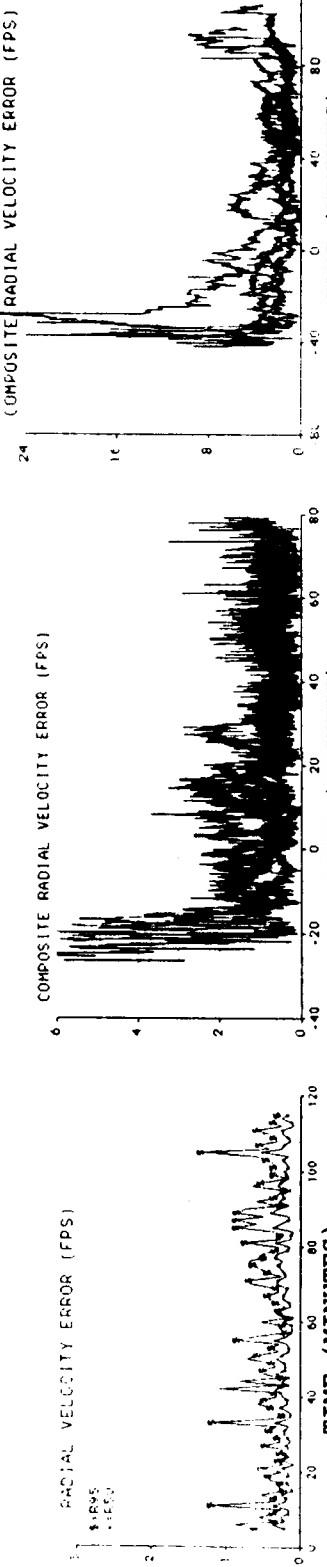


Figure 9
Full-Up Mode, Sonobuoy Profile
(50th and 95th Percentiles)

Figure 11
INS/DVS Mode, Over Land

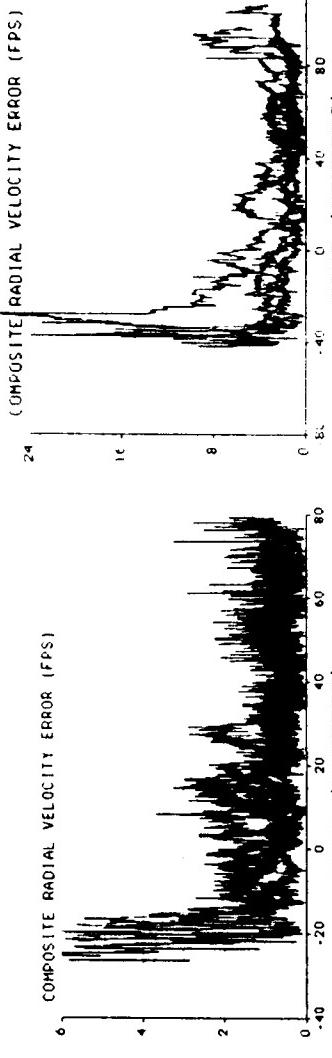


Figure 13
INS/DVS Mode, Over Water

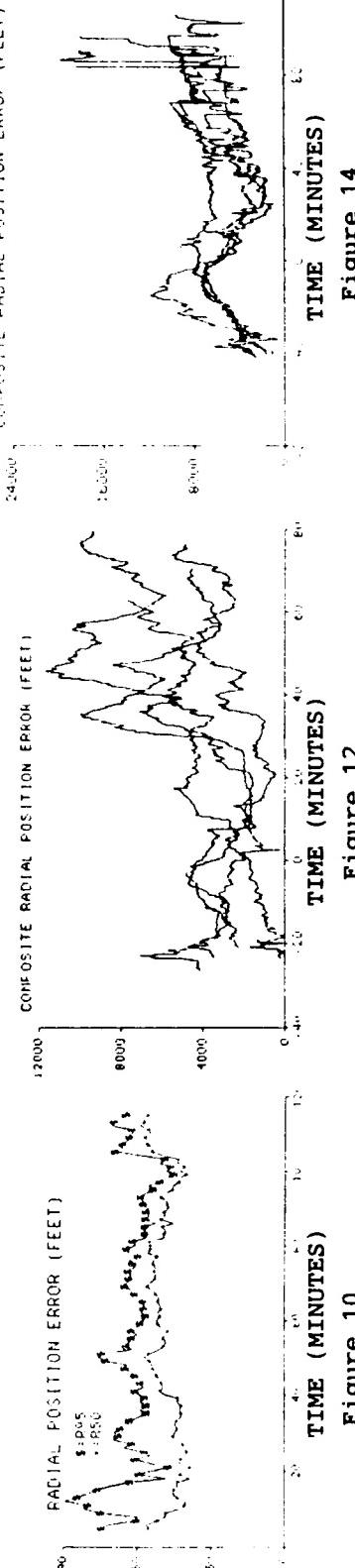


Figure 14
INS/DVS Mode, Over Water

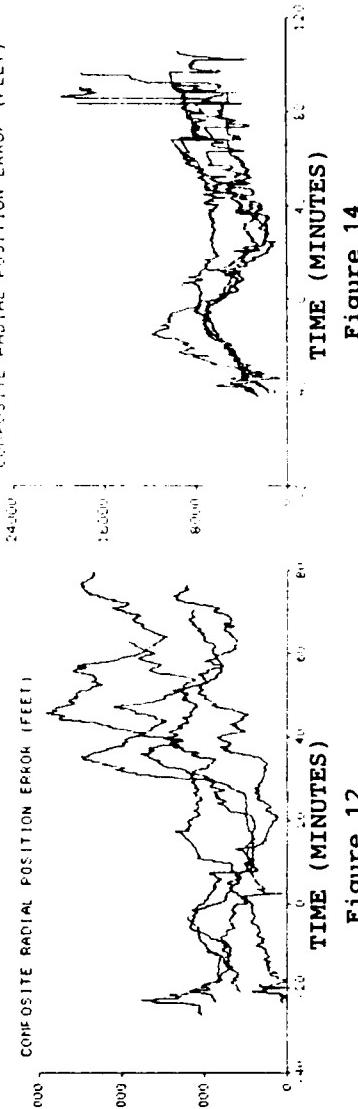


Figure 12
INS/DVS Mode, Over Land

Figure 10
Full-Up Mode, Sonobuoy Profile
(50th and 95th Percentiles)

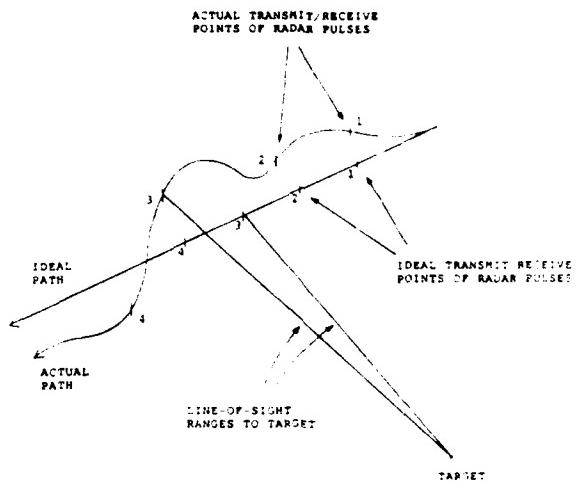


Figure 15
Spotlight SAR Geometry

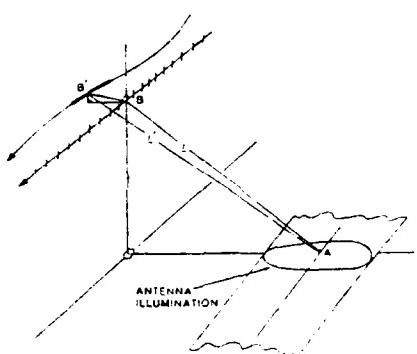


Figure 16
Stripmapping SAR Geometry

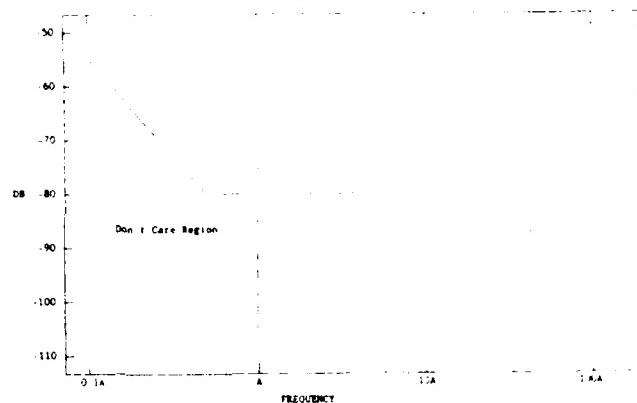


Figure 17 : Normalized Power Spectral Density of
Tolerable Antenna Phase Centre Displacement Error

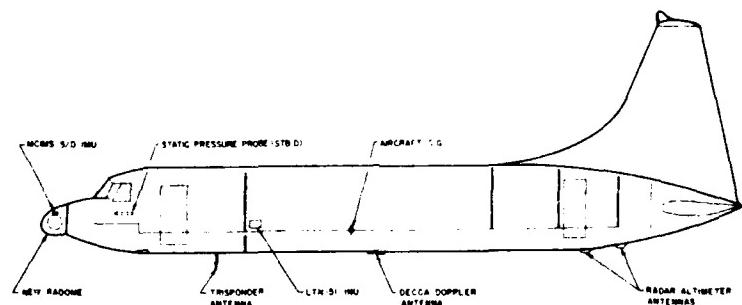


Figure 18 : Location of SARMC Instrumentation on
the NAE Convair 580

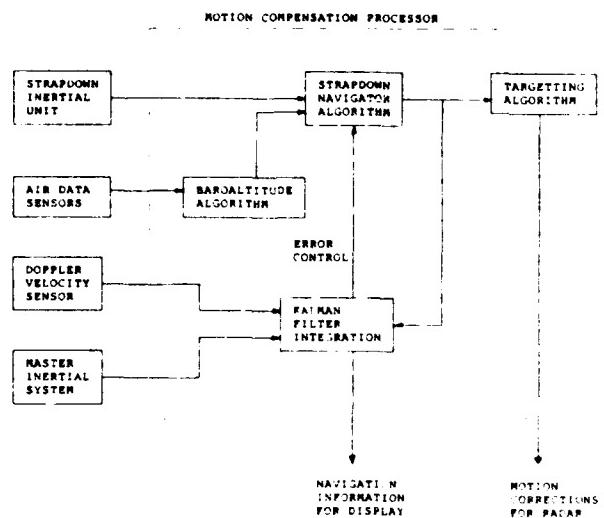


Figure 19 : Block Diagram of SARMC System

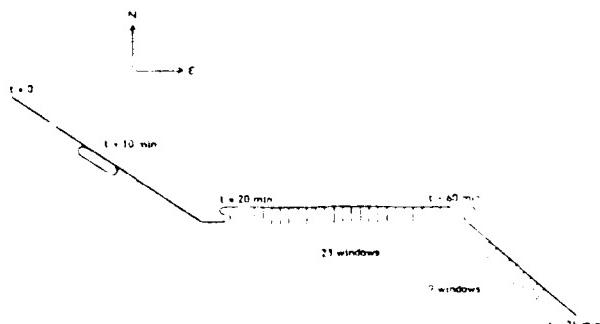


Figure 20 : SAR Mission Profile

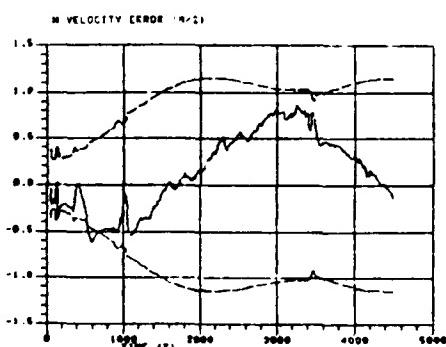


Figure 21
North Velocity Error & RMS For
Strapdown Navigator

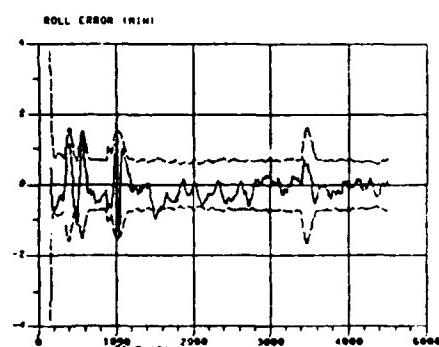


Figure 22
Roll Error & RMS For Strap-
down Navigator

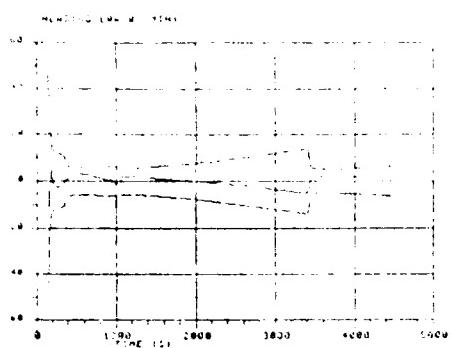


Figure 23
Heading Error & RMS For
Strapdown Navigator

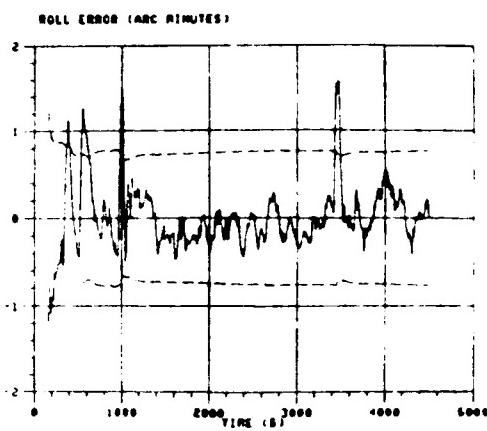


Figure 24
Roll Error and RMS for 21
State Kalman Filter

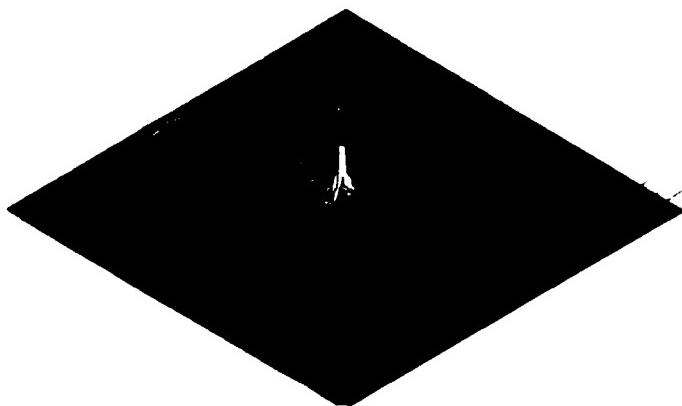


Figure 25 : Motion-Compensated Two-Dimensional
Impulse Response of the Radar

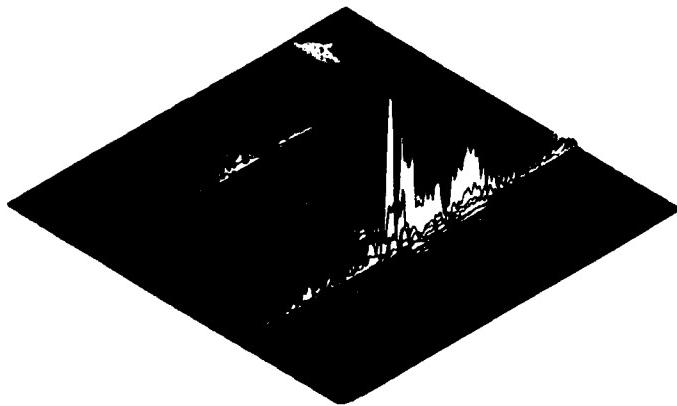


Figure 26 : Uncompensated Two-Dimensional Impulse
Response of the Radar



**Figure 27 : Motion-Compensated SAR
Stripmapping Image**

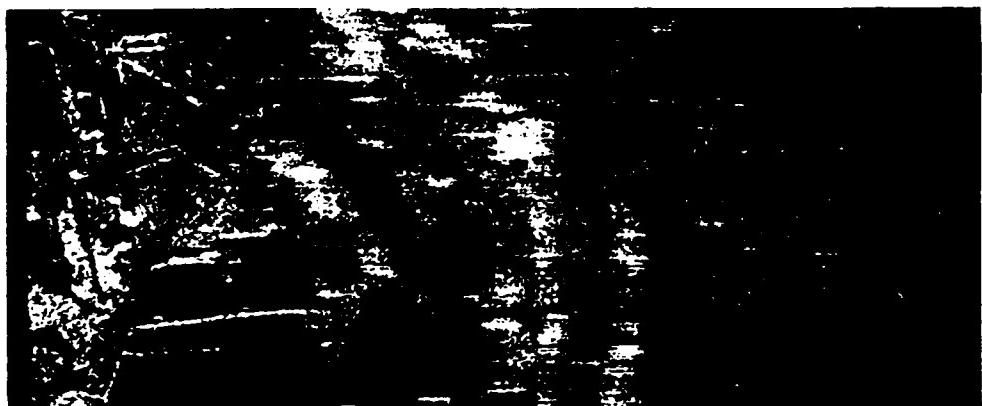


Figure 28 : Uncompensated SAR Stripmapping Image

OPPORTUNITIES AND CHALLENGES IN AVIONICS INTEGRATION, INS/GPS - A CASE STUDY

Zdzislaw H. "Stan" Lewantowicz, Colonel, USAF

Deputy Director, Avionics Directorate, Wright Laboratory

Adjunct Professor of Electrical Engineering, Air Force Institute of Technology

Danny W. Keen

Aeronautical Systems Division

Wright-Patterson AFB, OH 45433-6543

BIOGRAPHIES

Zdzislaw H. "Stan" Lewantowicz, Colonel, USAF, is the Deputy Director, Avionics, Wright Laboratory, Wright-Patterson AFB, OH. He served as Assistant Professor and Deputy Head, Department of Electrical and Computer Engineering, Air Force Institute of Technology, where he taught the integrated navigation system theory and applications courses. He has worked at the Air Force Weapons Laboratory in a high energy laser beam pointing, tracking, stabilization, and wavefront control. Colonel Lewantowicz is a Senior Pilot with over 2500 flying hours. He received the B.S.E.E. from Lowell Technological Institute (currently University of Lowell) in 1968, the M.S.E.E. from AFIT in 1976, and the M.S.A.A. from MIT in 1986. His research interests are in integrated navigation systems, flight control, adaptive optics, and system modelling and analysis.

Mr. Danny Keen is a graduate of Wright State University, where he received his Bachelors Degree and Masters Degree in Systems Engineering in 1977 and 1980, respectively. From 1977 till the present Mr. Keen has worked for Aeronautical Systems Division (ASD) at Wright Patterson AFB, Dayton Ohio, primarily in the navigation/guidance area. Mr. Keen has provided support in varying capacities to numerous ASD programs including the A-10, B-1, ACM, B-2, and C-17. Mr. Keen has also worked as chief project and flight test engineer in the 4950TH Test Wing at Wright Patterson AFB. Mr. Keen currently serves as the Technical Specialist for the ASD Avionics Engineering Navigation Branch. Mr. Keen's previous publications include papers on various navigation/flight test topics for NAECON, the American Defense Preparedness Association, and NATO AGARD panels.

ABSTRACT

Several recent technological developments (including the advent of GPS and the virtual explosion of computational capability) offer unprecedented opportunities for radically different approaches to providing weapon system capabilities. In the past high-rate raw sensor data would have to be sampled, or significantly pre-processed in a separate (dedicated) computer, before such data could be used within the constraints of a real time weapon system application. During such sampling/pre-processing much of the information content could be lost. The current computational capability explosion now permits highly complex algorithmic solutions to be executed within real time. Such computational capabilities within an integrated avionics architecture offer opportunities for maximum rate fusion of all raw signals directly from the sensors, where they are generated, with little or no information loss. As an example, when unfiltered GPS information is directly/optimally combined at maximum rate with inertial data, not only is the navigation accuracy increased, but effective susceptibility to GPS signal dropout/jamming is significantly reduced. Additionally, the resultant highly precise GPS/INS position, velocity, attitude, and time information can serve as a common integrating agent for combining the information from a variety of other mission sensors and communication devices. A system engineering process is required to

systematically explore and exploit the yet untapped opportunities presented by these technological advancements. The essential ingredients of this process are a system level perspective, sound modeling and simulation, and the application of analytic tools.

INTRODUCTION

The potential of GPS to contribute to the military mission is no longer in question. The acceptance by the USAF Major Commands was already well established approaching the end of 1990. Former commander of the Tactical Air Command (TAC), General Russ, stated during the *TAC Day* in July 1990 (in essence) "GPS is here to stay, we have tried it, our pilots love it, get it in our aircraft as quickly as possible." The Operation Desert Storm simply validated and amplified this acceptance. Numerous articles in the civil, technical, and military journals have expounded on the GPS contribution to that operation. One defense journal, *Defense News*, quoted senior Defense officials that of the numerous success stories of the various space system asset contributions to that desert war, that of GPS was most significant. In the Global View section of the June 1991 edition of the *GPS World*, in an article entitled "After the Storm: DoD Officials Back Selective Availability," USAF General Donald J. Kutyna, Commander-in-Chief of the U.S. Space Command, is quoted: (it) "would be difficult to overstate the value of the NAVSTAR GPS system during Desert Shield and Desert Storm operations" and "this system has proved to be an unqualified success."

In Operations Desert Shield and Desert Storm the Allied forces also demonstrated tremendous ingenuity in the use and application of the available GPS equipment. Although the GPS receiver equipment available to them performed very well, the system has the potential for providing significantly greater capability. Had the full GPS system potential been more flexibly exploited, the Desert Storm air campaign would have likely produced even more impressive results.

The challenge for the military research and development community is to vigorously exploit the simultaneous arrival of the GPS, the explosion in computational capability, and availability of the integrated avionics architectures for weapon systems. These factors offer unprecedented opportunities for much greater exploitation of avionics sensor fusion. The level of sensor fusion achievable is strongly influenced by the avionics architecture. The avionics architecture can also significantly enhance or degrade the sensor fusion results.

FEDERATED ARCHITECTURES

Historically, the aircraft avionics have evolved from dedicated, single function, mechanical sensor systems to, more recently, sensor systems which have become quite sophisticated in their functionality and accuracy. The earlier sensors were developed, refined, and added to the aircraft as stand-alone (federated) devices which provided crew members the increased ability to perform their mission more effectively. The fusion of information from a multitude of these avionics subsystems was a required pilot (or crew member) skill. For example, instrument flying requires the

pilot to methodically and rapidly "cross-check" several instruments that make up the "performance" and "control" instrument groups. He must absorb this multitude of information and make continuous and real-time decisions about the type and quantity of control inputs to apply in order to achieve the required flight path precision. In this case the pilot mentally performs the fusion of the various elements of information required for safe instrument flight. Incidentally the auto-pilot, when engaged, performs similar fusion and control decisions for a subset of the flying tasks.

Increases in the earlier avionics capabilities typically came from evolutionary increases in sensor accuracy, range, reliability, affordability, etc. The more recently developed sensors have embedded in them quite impressive computer processing capabilities to enhance sensor performance. In more recently developed sensors, functional outputs from other sensors are added to enhance their performance. However the provided information, being the output of another sensor, is typically processed rather than the unprocessed information available closer to the sensor detector. In the past there had been valid and practical reasons for this. Two of these reasons were the technology availability and standardization constraints.

In recent past, computational and data bus technologies either did not exist, or their performance was inadequate, or the mission did not require the "complexity." With time, as technology advanced, the sensor performance increased, the computational and data bus capabilities were developed and improved, and the mission requirements "warranted" some communication of various outputs among the avionics black boxes. However, each sensor evolved primarily from the necessity for stand-alone operation, that is without the necessity for communication of much information to or from other sensors. Where information was needed to be shared with other sensors, the format of that information was typically constrained to that available at the post-processed output serving the primary sensor purpose.

Despite some challenges in providing for multiple requirements, this limited integration concept was successful in that performance of some sensors was improved, to a lesser or greater degree, by adding to it information from other sensors. Most of the information that was provided from other sensors was at the post-processed output level or in the domain of output integration. This post-processing modifies the "raw" signal, available from the detector, in bandwidth, noise statistics, and other electronic compensation characteristics. Although theoretically it is possible to "undo" the output processing, if the original processing is known, this is not practical due signal-to-noise losses, distortions, and typically high cost. The impact of this post-processing is that it limits the degree to which the information from other sensors can be utilized, where appropriate, to improve the performance of other sensors.

ADVANCED AVIONICS ARCHITECTURES

In contrast to the federated avionics architectures, the integrated avionics architectures offer numerous advantages. These advantages can be understood from a broader perspective of avionics system flexibility, redundancy, fault tolerance, real-time reconfigurability, decreased life cycle cost, adaptability, and other attributes. These architectures consist of multi-function, software programmable modules in common avionics enclosures. Many of the mission functions can be performed by individual or groups of modules. This concept replaces the discrete functions of the federated avionics architectures. The Integrated Communications, Navigation, Identification Avionics (ICNIA) and the Integrated Electronic Warfare System (INEWS), are examples of such architectures, and are the basis for the U.S. Air Force Advanced Tactical Fighter (ATF), recently designated as the F-22, and the U.S. Army's next generation helicopter the LHX. The U.S. Navy is considering this avionics architecture for its AX system.

Selection of modular avionics standard is a topic of high interest to the North Atlantic Treaty Organization (NATO), in particular the Four Power Group consisting of United States, United Kingdom, France, and Germany. The essence of the question before this group is not whether integrated avionics should be considered, but what form should the standards take for the avionics architectures and for one of its key elements, the modules.

Many advantages result from the integrated architecture concept. The essentially raw signals are brought into a single framework where real-time management of the available resources can be exercised. This immediately offers opportunities for real time reconfiguration for a spectrum of requirements, be it to account for failures, or reconfigure to adapt to a changing mission phase, or to respond to changing mission requirements, such as a just identified enemy threat dictating a unique and immediate response.

Perhaps the most significant advantage of integrated avionics, as the main message of this paper, is that most of the signals are available in various forms and stages of processing within the physical confines of this architecture. Sensor fusion ranging from combining of two or three sensors to large scale fusion of a multitude of sensors is possible. For example, availability of the "raw" (unprocessed) pseudo- and delta-range GPS measurements provides unprecedented opportunities. When the GPS is properly combined with INS, this two-sensor subset produces a significantly more accurate, robust, jam-resistant, less costly navigation, attitude, and time reference set. This reference set is very important because not only does it provide a superior navigation solution, but it also becomes an integration reference set, for combining (fusing) of information from other sensors.

When other sensor information, which contains either absolute or relative navigation, attitude, or time information, is processed jointly with that of GPS and INS, then additional benefits are available. For example the reference biases, and other errors, of these sensors are often observable and estimable, thus providing a cross-calibration function. For those sensors where this cross-calibration is significant, several opportunities immediately appear. For example, the absolute performance tolerances on those sensors perhaps could be relaxed, thus reducing the complexity, cost, and reliability of those sensors. In some cases more than one sensor performs a similar or an overlapping function, or measures information also measured by some other sensor. When the information from this ensemble of sensors is properly fused, perhaps some of those sensors are either redundant, or their omission could result in minor and acceptable mission performance degradation but at a significant reduction in cost. Another possibility may be the realization of a common and highly accurate time reference that could be derived from this joint fusion of signals which have the GPS time as a common reference. Also, when information from other sensors is jointly fused with that of GPS and INS, the reference set accuracy and robustness is further improved.

These are only few examples of what may be achievable with integrated architectures. It is essential to emphasize, that to realize this vision, system and component level modeling, analysis, and simulation are required to identify the system level sensitivities and perform various trade-offs. The following sections of this paper describe a GPS/INS integrated avionics architecture case study that was first documented in Reference 1.

CONCERN WITH CURRENT MILITARY GPS/INS INTEGRATION APPROACHES

The federated architecture, with its limitations, is the prevalent environment that current military GPS user equipment is designed to operate within. Neither the Standard GPS Receiver IIIA nor the new Miniaturized GPS Airborne Receiver (MAGR) provide raw (ionospheric delay corrected) pseudo-range and delta-range data from the GPS satellites. Instead these receivers pre-process this

data and output position and velocity data as outputs of their own internal Kalman filters. These filters are based on a generic INS error model that is not optimized to any particular INS type or technology. If an aircraft requires a navigation solution optimized to a particular INS or if other sensors are required in the optimal navigation solution these GPS receivers constrain the aircraft Kalman filter designer to a cascaded filter approach. Use of GPS filtered position and velocity outputs to drive a separate aircraft Kalman filter can lead to filter stability problems, since the GPS and the INS position and velocity data are time correlated (References 2 and 3). This potential problem is minimized by typically processing GPS measurements in the aircraft filter at a much slower rate than they are available from the GPS receiver. This helps in reducing the correlation between GPS and INS measurements. A spacing of 10-12 seconds is sufficient when 4 satellites with good geometry are available. However, when fewer than 4 satellites are available, the GPS receiver filter outputs degrade and closely track the rapidly growing INS errors. To avoid filter instability under these conditions the aircraft filter is designed to disregard inputs from the GPS receiver when fewer than 4 acceptable satellites are available. This cascaded filter approach therefore results in the binary on/off use of GPS information.

The Standard GPS Receiver IIIA and MAGR could easily be modified to support a more integrated avionics architecture. If pseudo-range and delta-range data were directly output from the Standard GPS Phase IIIA Receiver or the MAGR, a non-cascaded aircraft Kalman filter could easily be implemented within the computational capability of today's avionics computers. Such a filter could use 100 percent of the available raw data to provide graceful position and velocity performance degradation when fewer than 4 satellites are available. In the future more complete use of available data will be required. Many current and future aircraft will be built as "low observable (LO)" vehicles. The addition of any aperture negatively impacts LO characteristics of such antennae. To minimize the LO impact GPS antenna gains will have to be sacrificed with a corresponding reduction in GPS performance. Up to 10 dB of reduced antenna gain at low scan angles is likely. Another concern for GPS signal availability is during, or near the weapon delivery time when enemy jamming is likely to be present. Even use of momentary and random glimpses at individual satellites by a non-cascaded filter, during a high jamming period should provide significantly improved performance over the cascaded filter approach (Reference 4). Following sections of this paper describe the system engineering process performed to define an integrated avionics approach for implementing GPS/INS non-cascaded Kalman filters. Potential sets of error states required for such non-cascaded filters are determined. Subsequently, performance of these filters is quantified via analysis against various potential GPS coverage conditions.

PROPOSED GPS/INS INTEGRATED AVIONICS APPROACH

Modify the GPS receiver IIIA and the MAGR to output pseudo-range and delta range, satellite position, and GPS time tag data on the MIL-STD 1553 data bus. Develop a joint Kalman filter that uses these GPS pseudo-range and delta-range measurements (instead of the filtered GPS position and velocity outputs). Include as a minimum the addition of one and/or two GPS receiver clock bias and drift states to this joint Kalman filter. Model the effect of additional GPS line-of-sight (LOS) measurement errors as states in the Kalman filter (should result in higher estimation accuracy). Furthermore, since raw GPS measurement errors are not time correlated with those of the IMU, all GPS data can be used at the maximum GPS output rate. This enables IMU errors, receiver clock bias and drift, and other errors to be rapidly and accurately computed in the aircraft integrated (joint in the probability density function sense) Kalman filter. If, at a subsequent time, fewer than 4 satellites are available the now calibrated receiver clock and

other GPS and IMU errors will facilitate use of pseudo-range and delta-range data from the available satellites for at least 30 minutes to keep the navigation solution bounded (Reference 5).

SIMULATION TOOLS

To quantify the level of performance improvement expected by the above proposed GPS/INS integrated avionics approach, the following simulation tools are used:

a) Multimode Simulation for Optimal Filter Evaluation (MSOFE), a Fortran based program capable of multimode simulations, "assists the engineer in the design and evaluation of integrated systems that employ Kalman filtering techniques" (Reference 6). MSOFE was designed jointly by Integrity Systems, Inc., and the (former) Avionics Laboratory, WRDC/AARN, (currently WI/AAAS is a branch in the Avionics Directorate of the Wright Laboratory) Wright Patterson AFB. The greatest advantage in using MSOFE is its time and cost saving benefits. By utilizing its 14 problem-specific subroutines and its 59 general subroutines, a user may perform evaluation on a system in 20% of the time previously required. It is referred to as multimode in that it is capable of simulating a variety of system problems. Its two most widely used modes are Monte Carlo and Covariance simulations.

b) Profile generator (PROFGEN) is a trajectory generation program developed by the (former) Avionics Laboratory at Wright-Patterson AFB. It computes position, velocity, acceleration, attitude, and attitude rates for a maneuvering air vehicle profile based on user provided inputs. This flight profile is directly usable by MSOFE. A detailed description of PROFGEN is found in Reference 7.

FILTER DEVELOPMENT

A 98-error state truth model was developed for the fully integrated GPS/INS navigation system. It combines 68-state IMU truth model based on the LN-94, the Litton Inc. version of the Air Force Standard RLG INS (Reference 8) and a 30-state GPS truth model developed by Solomon in support of an advanced Completely Integrated Reference Instrumentation System (CIRIS) for the Central Inertial Guidance Test Facility (CIGTF), Holloman AFB, NM (Reference 9). From this truth model various reduced order filters are created and evaluated using MSOFE against the representative mission trajectory. The performance of each of these reduced-order filters is discussed in the Filter Evaluation Section herein.

The first reduced order (51-state) filter is created by identifying and deleting those error states whose error contributions are multiple orders of magnitude smaller than the more dominant error states. Additionally other error states, with similar coupling paths, are linearly combined. Fifteen IMU states are one or more orders of magnitude smaller in error contribution than other significant error states effecting either gyro platform tilt or velocity accuracy (e.g., the gyro scale factor asymmetry errors). Also 12 other minor IMU error states are sufficiently similar to other existing error states that they can logically be combined (e.g., gyro misalignment errors with gyro scale factor error). Similarly the twelve GPS satellite position error states are not fully observable, except for the position error vector projection onto the respective line-of-sight, and can be deleted or combined. Additionally, the ionospheric and satellite clock errors are combined with the tropospheric errors to reduce the GPS model by another eight states. Except for the receiver clock errors, the GPS measurement errors to each satellite are assumed to be statistically independent. (In fact some modest correlation does exist. For example, the low altitude atmospheric delay has some correlation as a function of angular proximity of the lines-of-sight and elevation angles.) Using this approach the 98-state truth model is reduced to a 51-state reduced-order filter. These 51 states are listed in the Appendix. The resulting 51-state reduced-order filter is tuned and its performance compared to that of the truth model. As expected little or no tuning noise is added to selected states to

insure filter stability (filter tuning).

The next 32-state reduced-order filter combines the effects of the following 19 states into the remaining error terms (see Appendix).

- a. 3 Gravity Anomaly States
- b. 3 Accelerometer Correlated Bias States
- c. 3 Velocity Scale-Factor Asymmetry States
- d. 6 Accelerometer Misalignment States
- e. 4 GPS Code Loop Error States

These 19 error states are primarily the most insignificant error contributors of the remaining 51 states. The resulting 32-state reduced-order filter is tuned and its performance compared to that of the truth model. The performance of the 32-state reduced-order filter is slightly degraded from that of the 51-state filter.

The subsequent, 23-state reduced-order filter is generated by combining the effects of the following 9 states into the remaining error terms (See Appendix).

- a. 3 Accelerometer Scale Factor States
- b. 3 Gyro Scale Factor States
- c. 3 Vertical Channel States

These 9 error states are the most insignificant error contributors of the remaining 32 states. The resulting 23-state reduced-order error filter is tuned and its performance compared to that of the truth model. A moderate amount of tuning noise is added to insure filter stability. As expected the performance of the 23-state reduced-order filter is somewhat degraded from that of the 32 and 51-state filters.

The final 18-state reduced-order filter is generated by deleting the following 5 states (see Appendix).

- a. 4 Code Loop States
- b. User Clock Drift

These 5 error states are the least significant of the 23 primary error contributors. Deletion of these states should only be considered under severe computational limitations. A significant amount of tuning noise must be added to insure filter stability, especially for those conditions where fewer than 4 satellites are available. The performance of the tuned 18-state reduced order filter is significantly poorer than of the 51, 32 or 23-state filters. However, the projected performance of the joint 18-state reduced-order filter promises to be better than the F-16 15 state cascaded filter performance demonstrated during recent F-16 GPS/INS flight testing (Reference 10). With four satellites available, the projected performance of the joint 18-state filter is 10 feet compared to the demonstrated performance of 27 feet of the F-16 cascaded filter.

FILTER EVALUATION

Before the simulated flights each filter is aligned using the flight filter. No separate alignment filter is required. In fact aligning with the flight filter has several advantages. During alignment zero velocity, barometric altimeter, and the four GPS measurements are incorporated. This provides the filter with the opportunity to immediately begin estimating those error states which are observable. The significant advantage during alignment is that the flight filter begins building up the cross-correlation terms in the covariance matrix between those observable states.

A representative attack aircraft weapon delivery mission is generated using PROFGEN. This mission flight trajectory is shown in Figure 1. The top portion of the Figure shows the entire mission profile. The significant segments during the 0 - 2000 second period of the simulation are the takeoff, cruise, and descent. The significant segments during the 2000 - 2800 second period are the ingress, pop-up maneuver, weapon delivery, evasive maneuvers, and escape. This period is shown as an enlargement in the bottom portion of Figure 1. This is the segment that is

examined with various GPS satellite availabilities. The final period, 2800 - 7295 seconds, is the return of the aircraft to the same base. Single-run Monte Carlo simulation is performed for each case analyzed for relative filter performance comparisons.

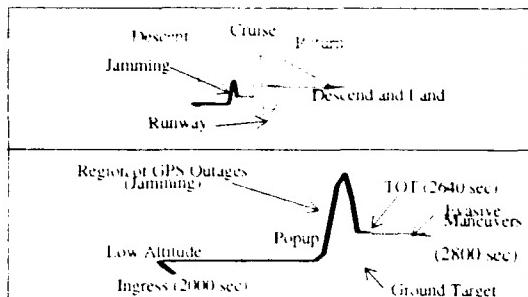


Figure 1. Flight Profile

The reduced order filters described above are exercised using MSOFE simulation against this mission profile. During the take-off, climb out, and cruise enroute to the target area four GPS satellites are simulated as continuously available until the 2000 second (33.33 minute) point. A single-run Monte Carlo simulation is performed from take-off until the 2000 second point for each reduced order filter and the truth model. This occurs immediately after the completion of descent to a low altitude.

The position and velocity error performance plots, for this initial 2000 second segment, are shown in Figures 2a and 2b. The appropriately labeled performance plots provide the Kalman filter confidences expressed in 1-sigma values for the tuned 18, 23, 32 and 51-state reduced order filters (gray line) as well as the estimation of error (black line). These are presented as root sum squared (RSS) three-dimensional position and velocity errors (magnitudes of the position and velocity error vectors) versus time. No attempt is made to locally optimize the performance of any of the filters.

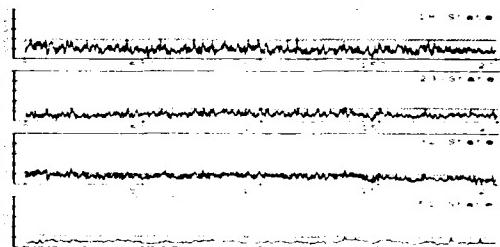


Figure 2a. 3-D Position Error (ft), Take-off to 2000 Sec.

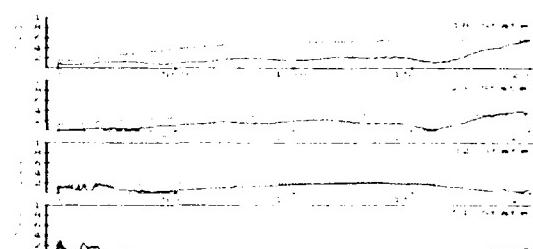


Figure 2b. 3-D Velocity Error (fps), Take-off to 2000 Sec.

At the 2000 second point into the mission, and approximately 10 minutes prior to time-on-target (TOT), varying qualities of GPS

coverage are assumed. A separate Monte Carlo simulation run is performed for an additional 800 second (13.33 minute) period for each reduced order filter against each assumed GPS coverage case. Each GPS coverage case, and its associated position and velocity estimation performance plots for each reduced order filter is illustrated in Figures 3a and 3b through Figures 8a and 8b below. These plots provide the Kalman filter confidences (1-sigma values) and nominal estimation errors as three-dimensional RSS position and velocity error plots.

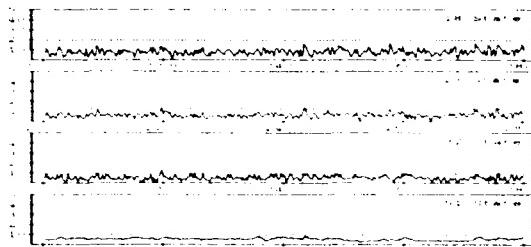


Figure 3a. 3-D Position Error (ft), Four Satellites

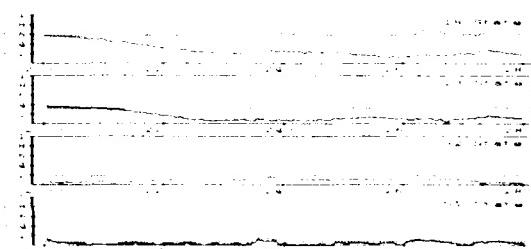


Figure 3b. 3-D Velocity Error (fps), Four Satellites

The four-satellite performance is the baseline for comparison of the following three, two, one, zero and occasional satellite measurement availability cases. The GPS satellite geometry assumed in this analysis is with a GDOP of approximately 3 with one satellite directly overhead. Note the relative performance of the more significantly reduced-order filters against the 51-state filter, whose performance is very close to that of the truth model (98-state filter) performance (not shown). Also note the superior performance of the 51-state filter, especially in velocity. When the four satellite measurements are available, the position performance differences are not significant, approximately 6-7 ft for the 51-state case compared to approximately 8-10 ft for the 18-state case. The velocity performance is significantly better for the 51-state versus the 18-state case.

When the overhead satellite is assumed not available (the three-satellite case), the 18-state filter position estimation error performance (black line) degrades by approximately a factor of two. The performance of the other reduced order filters remain essentially the same as the full satellite coverage case. Note that the 18-state Kalman filter 1-sigma position confidence plot (gray line) is significantly larger than that of the other filters. This large difference in filter confidence is driven by the moderate amount of tuning noise incorporated in the 18-state filter to insure filter stability under conditions of little or no satellite availability. Without this additional tuning noise the 18-state filter became unstable.

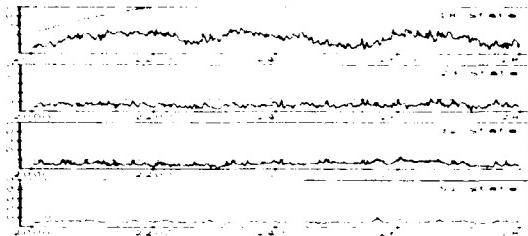


Figure 4a. 3-D Position Error (ft), Three Satellites

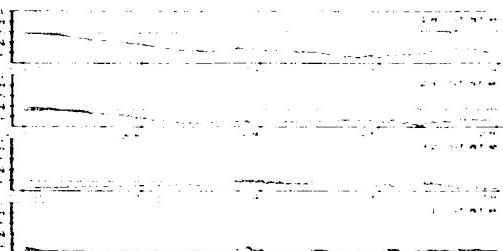


Figure 4b. 3-D Velocity Error (fps), Three Satellites

In the next case analyzed, in addition to the loss of the overhead satellite, one more satellite is removed leaving measurements from only two satellites available to the reduced order filters during the ingress and attack portion of the mission. Note that in all of the cases analyzed, each reduced order filter includes at least one state for the barometric altimeter, and the associated atmospheric pressure variation errors which are modeled as first-order Markov random processes. The presence of this state appears to significantly enhance the performance of each filter, including the 18-state filter. In the two-satellite case the position error performance of the reduced order filters varies substantially. The 32 and 51-state filters perform significantly better than the 23 and 18-state filters. This is especially true during the first 400 seconds (or so) of this flight segment.

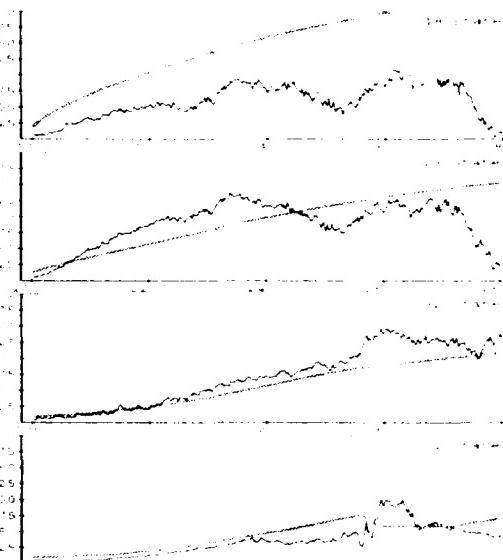


Figure 5a. 3-D Position Error (ft), Two Satellites

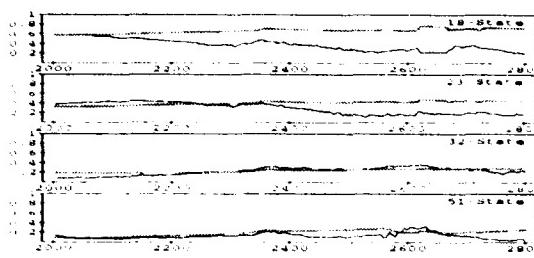


Figure 5b. 3-D Velocity Error (fps), Two Satellites

The next case analyzed is that of a single satellite and barometric altimeter measurements. The position error performance is somewhat degraded in comparison to the earlier cases. Noteworthy is the velocity performance of the 51 and 32-state filters, where the error increases are minor. Also note the advantage of additional states in the two higher order filters, especially during the first 400 seconds.

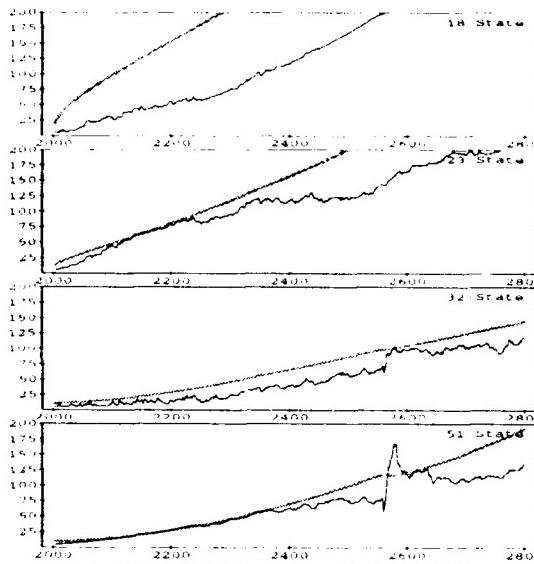


Figure 6a. 3-D Position Error (ft), One Satellite

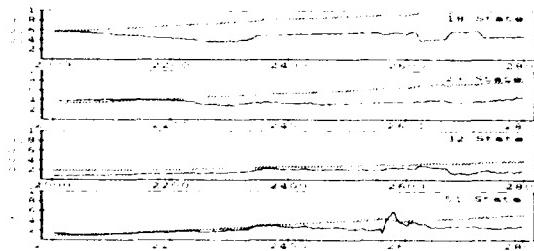


Figure 6b. 3-D Velocity Error (fps), One Satellite

Although a particular single satellite is chosen it is believed that very similar performance will result with any other single satellite, except the vertical one. The barometric altimeter is already providing substantial information in the vertical direction. The lone off-vertical satellite significantly helps bound the growth of horizontal position errors. The majority of the error growth observed in these position error plots is due to the loss of

estimation accuracy of the vertical channel (barometric) errors due to growing uncertainty of the atmospheric correlated Markov error growth.

Plots 7a and 7b show the performance of each filter when only barometric altimeter measurements are available. In this case each filter relies on the quality of estimates of the INS and barometric measurement errors. Of course the estimation quality of these errors depends on the estimation quality of the GPS measurement errors during the time when GPS measurements are available. The position error performance is surprisingly good for the 51-state filter compared to the other filters. The velocity error performance is also quite good, especially when compared to the velocity performance of cascaded filters (Reference 10).

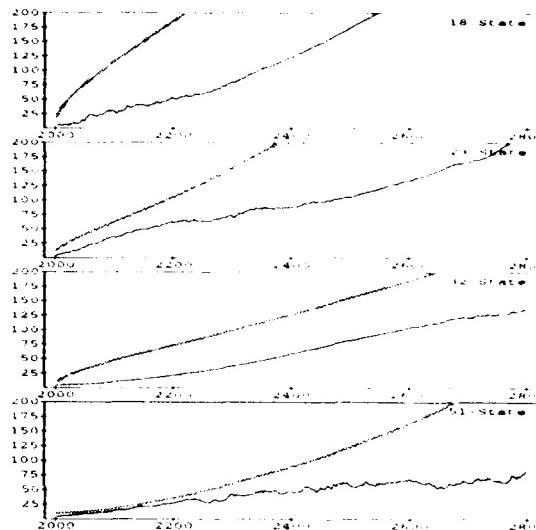


Figure 7a. 3-D Position Error (ft), Zero Satellites

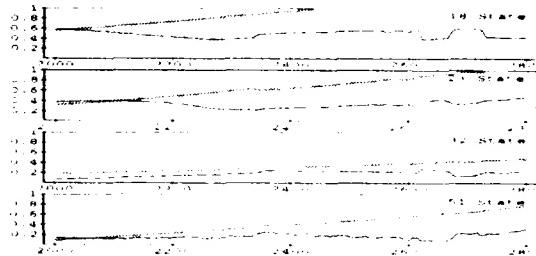


Figure 7b. 3-D Velocity Error (fps), Zero Satellites

This final case is the most interesting. It is a simulation of a situation where only occasional GPS measurements from various satellites are available. Although in reality the time distribution of the sporadic availability of GPS measurements is expected to be random, here it is approximated with a single satellite measurement available only once every 30 seconds on a rotating basis, that is each satellite is available only once every 120 seconds. The results are interesting, especially for the 32 and 51-state filters. The position error performance is near that of the four satellite case, within a factor of two or so. The velocity performance of each of these filters is nearly unchanged from the four satellite case.

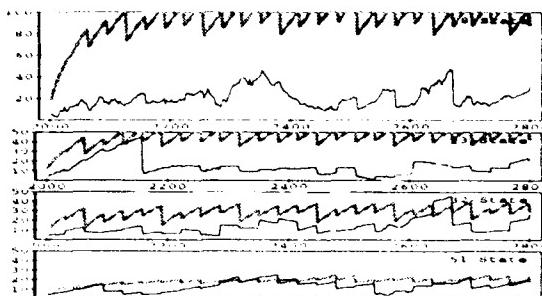


Figure 8a. 3-D Position Error (ft), Occasional Satellites

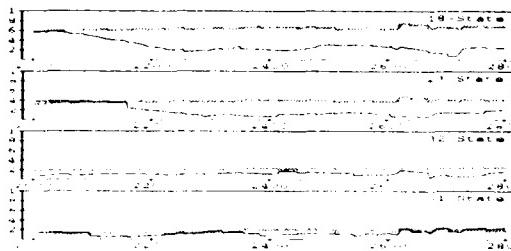


Figure 8b. 3-D Velocity Error (fps), Occasional Satellites

SIMULATION RESULTS

The RSS position and velocity error results for each assumed GPS coverage case are shown in the previous Figures 3a - 8b. These simulation results predict that a non-cascaded Kalman filter receiving pseudo-range and delta-range measurements and modeling significant GPS and barometric altimeter measurement errors will perform significantly better in actual flight than the current cascaded approach.

The results of the zero satellite case provide a "best case" approximation (upper bound) for cascaded filter performance, with fewer than 4 satellites under these same conditions. This is because the cascaded filter does not use the filtered GPS receiver computed position and velocity data when fewer than 4 satellites are available. Note that for all the reduced order filters the RSS performance at the 10 minute point for the 3-satellite case is essentially the same as for the 4 satellite case, with the exception of the 18-state filter. For the one, two, and three satellite cases the 23-state RSS position error values at the 10 minute point are 130, 100, and 10 respectively. These values are significantly better than the "best case" value that could be expected for the cascaded filter approach using fewer than 4 satellites. Similarly the 23-state velocity errors at the ten-minute point for the one, two and three-satellite cases are 1.0, 0.8, and 0.6 feet per second. Note the error growth rate for all Kalman filters is substantially less for the 1 satellite case versus the zero satellite case. Another interesting case is when only occasional single GPS measurements are simulated by a single measurement from each of four satellites, in turn, at 30 second intervals. It is clear for all the reduced order Kalman filters that even occasional pseudo-range measurements provide sufficient information to significantly improve filter performance over that of the cascaded filter case.

CONCLUSIONS AND RECOMMENDATIONS

It is clear that the choice of avionics integration architecture can inherently enhance, enable, or severely limit potential mission capabilities. This choice of integration architecture, can also directly and profoundly affect performance, cost of integration, and cost of ownership. These impacts are clearly illustrated by the GPS/INS case study.

The above GPS/INS case study demonstrates that a significant level of navigational performance improvement can be realized for the case where fewer than four satellites are available by modifying the standard GPS IIIA receiver and MAGR to output pseudo-range and delta-range data on the MIL STD 1553 data bus and by incorporating a non-cascaded integrated aircraft Kalman filter.

The quality of error estimates improves significantly during any maneuver segments, be it horizontal or vertical. In fact the flight trajectory is rather sterile in the sense that all maneuvers are perfect. A straight and level flight is precisely straight and level. In reality pilots, and to a lesser degree autopilots, are not physically capable in flying these segments as perfectly. Thus any rolling, pitching, yawing, or linear accelerations, even relatively small in magnitude should help the filter by making the less observable errors become more observable.

Additional significant performance improvement can be realized by incorporating as many significant error states into this single integrated Kalman filter as possible based on the aircraft memory and the throughput constraints.

For the applications where the position, velocity, and attitude accuracy are the primary system contributors to weapon-on-target miss distances, this GPS/INS integration approach promises to significantly reduce those error components and provide significantly more robust mission capability.

REFERENCES

1. Lewantowicz, Zdzislaw H., and Keen, Danny W., "Graceful Degradation of GPS/INS Performance With Fewer than Four Satellites," Proceedings, Institute of Navigation, National Technical Meeting, January 1991.
2. Cunningham, J. R., and Lewantowicz, Z. H., "Dynamic Interaction of Separate INS/GPS Kalman Filters," Proceedings of the The Institute of Navigation GPS 88 International Symposium, September 1988.
3. Johnson, G. B., and Lewantowicz, Z. H., "Closed Loop Operation of GPS Aided INS," Proceedings of the Institute of Navigation GPS-90 International Symposium, September 1990.
4. Lewantowicz, Zdzislaw H., and Keen, Danny W., "GPS Aided Navigation Performance for Weapon Release With Fewer Than Four Satellites," Proceedings, Joint Services Data Exchange Conference, October 1990.
5. Hirt, E. H., and Yakos, M. D., "Time Dissemination Using NAVSTAR Global Positioning System (GPS) Phase IIB User Equipment," 35 Annual Frequency Control Symposium, USAERADCOM in Ft. Monmouth, NJ, May 1981.
6. Carlson, Neal, and Musick, Stanton, "User's Manual for a Multimode Simulation for Optimal Filter Evaluation (MSOFE)," AFWAL-TR-88-1138, Avionics Laboratory AFWAL/AARN-2, Wright Patterson AFB, OH, April 1990.
7. "PROFGEN - A Computer Program for Generating Flight Profiles," Avionics Laboratory AFWAL/AARN-2, Wright Patterson AFB, OH, November 1976.
8. "Performance Accuracy (Truth Model/Error Budget) Analysis for the LN-93 Inertial Navigation Unit," DID No. DI-S-21433 B/T: CDRL No. 1002, Litton Guidance and Control Systems, Woodland Hills, CA, January 1985.
9. Solomon, Joseph K., "CIRIS Special Study, EENG 699 Final Report," Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright Patterson AFB, OH, March 1989.
10. "Flight Test Report for Performance Validation of the Global Positioning System on the F-16 Block 40 Aircraft," General Dynamics Test Report No. 16PR8816, 19 June 1990

APPENDIX

State	Name	TM	F51	F32	F23	F18	State	Name	TM	F51	F32	F23	F18
INS													
$\delta\theta_x$	Angular Position	1	1	1	1	1	S_{QAX}	Accel. S.F. Asymmetry	51A	33			
$\delta\theta_y$		2	2	2	2	2	S_{QAY}		52A	34			
$\delta\theta_z$		3	3	3	3	3	S_{QAZ}		53A	35			
$\dot{\theta}_x$	Platform Tilts	4G	4	4	4	4	f_{xx}	Accel. Nonlinearities (9)	54A				
$\dot{\theta}_y$		5G	5	5	5	5	f_{yy}		55A				
$\dot{\theta}_z$		6G	6	6	6	6	f_{xz}		56A				
δV_x	Velocity	7A	7	7	7	7	f_{yx}		57A				
δV_y		8A	8	8	8	8	f_{yz}		59A				
δV_z		9A	9	9	9	9	f_{zx}		60A				
δh	INS Altitude	10H	10	10	10	10	f_{zy}		61A				
δh_l	INS Lagged Altitude	11H	11	11			μ_1	Accel. Misalignments (6)	62A				
δS_i	Vertical Channel Aid (K3)	12H	12	12			μ_2		63A	36			
δS_d	(K4)	13H	13	13			μ_3		64A	37			
b_{xc}	Gyro Correlated Bias	14G					σ_i		65A	38			
b_{yc}		15G					σ_2		66A	39			
b_{zc}		16G					σ_3		67A	40			
Δ_{xc}	Accel. Correlated Bias	17A	14						68A	41			
Δ_{yc}		18A	15										
Δ_{zc}		19A	16										
δg_x	Gravity Anomaly	20A	17										
δg_y		21A	18										
δg_z		22A	19										
δh_c	Baro Tropospheric	23H	20	14	11	11	δR_{clk}	User Clock Bias	69	42	27	18	18
δh_d	Baro Instrument Dynamics	24H					δD_{clk}	User Clock Drift	70	43	28	19	
δh_{sf}	Baro Scale Factor	25H					δR_{cdl}	Code Loop Errors (4)	71	44			
δh_b	Baro Fixed Bias	26H					δR_{d2}		72	45			
b_x	Gyro Bias	27G	21	15	12	12	δR_{d3}		73	46			
b_y		28G	22	16	13	13	δR_{d4}		74	47			
b_z		29G	23	17	14	14	δR_{trp1}	Tropospheric Delay (4)	75	48	29	20	
S_{gx}	Gyro Scale Factor	30G	24	18			δR_{trp2}		76	49	30	21	
S_{gy}		31G	25	19			δR_{trp3}		77	50	31	22	
S_{gz}		32G	26	20			δR_{trp4}		78	51	32	23	
X_1	Gyro Misalignment (6)	33G					δR_{ion1}	Ionospheric Residual (4)	79				
X_2		34G					δR_{ion2}		80				
X_3		35G					δR_{ion3}		81				
v_1		36G					δR_{ion4}		82				
v_2		37G					δR_{clk1}	SV Clock Bias (4)	83				
v_3		38G					δR_{clk2}		84				
D_{xxx}	Gyro Nonlinearities	39G					δR_{ik3}		85				
D_{yyy}		40G					δR_{ik4}		86				
D_{zzz}		41G					δx_{sv1}	SV Position Errors	87				
S_{Qbx}	Gyro S.F. Asymmetry	42G					δx_{sv2}		88				
S_{Qby}		43G					δx_{sv3}		89				
S_{Qbz}		44G					δv_{sv1}		90				
Δb_x	Accel. Bias	45A	27	21	15	15	δv_{sv2}		91				
Δb_y		46A	28	22	16	16	δv_{sv3}		92				
Δb_z		47A	29	23	17	17	δv_{sv4}		93				
S_{ax}	Accel Scale Factor	48A	30	24			δz_{sv1}		94				
S_{ay}		49A	31	25			δz_{sv2}		95				
S_{az}		50A	32	26			δz_{sv3}		96				
							δz_{sv4}		97				
									98				

AN OVERVIEW OF OPTICAL GYROSCOPES FOR NAVIGATION

J.G. Mark and D.A. Tazartes
 Litton Guidance and Control Systems Division
 5500 Canoga Avenue
 Woodland Hills, California 91367 USA

SUMMARY

In the 1980's, Ring Laser Gyrosopes (RLG) displaced the mechanical (spinning wheel) gyroscope as the angular sensor of choice for navigation. While the RLG remains the standard navigation grade instrument, several other optical gyroscopes have recently appeared. The multioscillator (or four-mode gyro) represents a new generation in laser gyroscopes. Systems based on this technology are now being delivered for use on commercial and military aircraft. Another optical sensor, the fiber optic gyroscope (FOG) has been incorporated in inertial measurement units (IMU) and proved itself capable of AHRS (attitude and heading reference system) accuracy. This gyroscope should find many applications in aided navigation systems. Integrated FOG/GPS systems appear attractive as low cost navigators. This paper addresses technology involved in these optical gyroscopes and discusses their advantages and disadvantages in relation to present and future applications.

LIST OF SYMBOLS

ac	- Alternating Current
AHRS	- Attitude and Heading Reference System
ccw	- Counter-Clockwise
cw	- Clockwise
deg	- Degrees
FOG	- Fiber Optic Gyroscope
GPS	- Global Positioning System
IFOG	- Interferometric Fiber Optic Gyroscope
IMU	- Inertial Measurement Unit
INS	- Inertial Navigation System
LCP	- Left Circularly Polarized
nmi/hr	- Nautical Mile per Hour
RCP	- Right Circularly Polarized
RLG	- Ring Laser Gyroscope
sec	- Second(s)

1. INTRODUCTION

Present-day optical gyroscopes employ the Sagnac effect¹ to permit observation of rotation relative to inertial space. Sagnac's work dates back to the early 1900's in conjunction with attempts to identify the "ether." However, the effect that bears his name and which deals with the behavior of counter-propagating light waves within a rotating cavity was later found to be rooted in general relativity. Application of this physical phenomenon to inertial sensors was not practical until the advent of the laser. In the early 1960's, laser gyroscopes were demonstrated for the first time and the interest in optical rotation sensors as replacements for the mechanical (spinning wheel) gyroscopes was launched. While the earliest embodiments of the laser gyro exhibited many problems, they also offered promise of improved reliability, extremely high dynamic range, and excellent linearity at high angular rates. The Sagnac effect also led to the FOG which, when properly mechanized, offers many of the same desirable properties with promise for lower cost. At present, there are two main classes of optical gyroscopes. The first type makes use of the Sagnac effect in resonant cavities.

It is comprised predominantly of several forms of RLGs. This category includes the conventional two-mode gyros as well as the multioscillator devices. The second type makes use of the Sagnac effect in an interferometer. The interferometric fiber optic gyro (IFOG) falls into this category. The discussion will focus on various forms of RLGs and the IFOG.

2. HISTORICAL AND MARKET PERSPECTIVE

Mechanical gyroscopes have long been used as rotation sensors for platform stabilization in guidance and navigation systems. While these devices demonstrated very good accuracy, the mechanical complexity associated with the spinning gyros and the stable platform led to compromises in reliability, weight, bandwidth, and dynamic range. With the advent of the laser, optical rotation sensors became a possibility and an intensive effort to develop the RLG was initiated in the early 1960's. The hope was to develop a relatively simple instrument whose accuracy would meet or exceed that achievable with the mechanical gyroscopes of the time but which would also offer better reliability and a potential for reduced cost. Another important consideration was that the wide dynamic range achievable with an optical gyroscope would permit the development of "strapdown" inertial navigation systems (INS) wherein the complex mechanical gimbals required to form a stabilized platform were eliminated. Indeed, the RLG used in a strapdown configuration has become the backbone of most present commercial and military navigation systems. Presently, the next generation of ring lasers is being developed and introduced into aircraft. New instruments include the multioscillator gyros which employ optical biasing to circumvent lock-in effects. Systems based on these gyros are being delivered for use in commercial and military aircraft. In addition, both the two-mode and multioscillator RLG systems have benefitted from improvements in component technology (particularly in the area of mirrors). This has resulted in smaller and lighter weight systems.

In parallel with RLG development, work was also being done in the field of fiber optics. The FOG was made possible by advances in the areas of light sources, low loss couplers, polarizers, and integrated optics. These components are important to the telecommunications industry but are also essential to fiber optic rate sensors. FOGs are attractive in that long fiber lengths can be used to obtain very fine angular rate resolution (the optical paths in FOGs are measured in hundreds of meters versus path lengths of tens of centimeters for RLGs). The ability to wind long lengths of fiber permits the construction of small gyros with acceptable performance.

The FOG was first considered to be a rate gyro but was later shown to possess rate integrating characteristics². In the area of navigation grade instruments, the RLG (both two-mode and multioscillator) remains clearly in the lead at the present time although navigation grade FOGs have been demonstrated and are included in future inertial navigation systems. On the lower accuracy end, the FOG promises many advantages, including a lack of moving parts, high angular rate resolution, low noise, low weight, cost, and power while retaining most of the

desirable characteristics of the RLG such as dynamic range and bandwidth. FOG-based IMUs have been delivered and demonstrated to meet their performance goals. These IMUs are very well suited to a wide variety of applications^{3,4} including smart munitions, rate gyros, missile guidance, and AHRS, as well as embedded GPS navigators.

3. SAGNAC EFFECT

The Sagnac effect deals with the propagation of light in a rotating reference system. Due to the rotation, the frame is not inertial and a general relativistic analysis applies. Interestingly enough, a classical treatment of the problem (as performed by Sagnac) yields the correct answer but for the wrong reason!

A simple view of the Sagnac effect consists of the following. Consider a closed path created by a cavity or waveguide in which counter-propagating (clockwise and counter-clockwise) beams are propagating. When the system is rotating, the effective optical path "seen" by the beam that is traveling in the same direction as the rotation becomes longer while the path "seen" by the beam traveling in the opposite direction becomes shorter. The difference in path lengths is a measure of rotation.

The Sagnac effect may be exploited for gyroscopic purposes in one of two different manners depending on whether the sensor is mechanized as a resonator or as an interferometer⁵. In a resonator (such as the various forms of RLGs), the clockwise (cw) and counter-clockwise (ccw) beams form resonant modes in a cavity. These modes can be viewed as a fictitious standing wave pattern which exists in the cavity. The Sagnac effect ensures that the standing wave remains fixed in inertial space and therefore provides a reference for measuring rotation. A detailed explanation of the RLG readout scheme appears in a later section.

In the case of an interferometer, light waves are launched in opposite directions into a cavity or waveguide. Rotation of this path causes an optical length difference between the two counter-propagating beams. This equates to an optical phase difference which can be detected by recombining the two beams as they exit the path. The interference phase observed is a measure of rotation.

4. RING LASER GYRO OPERATION

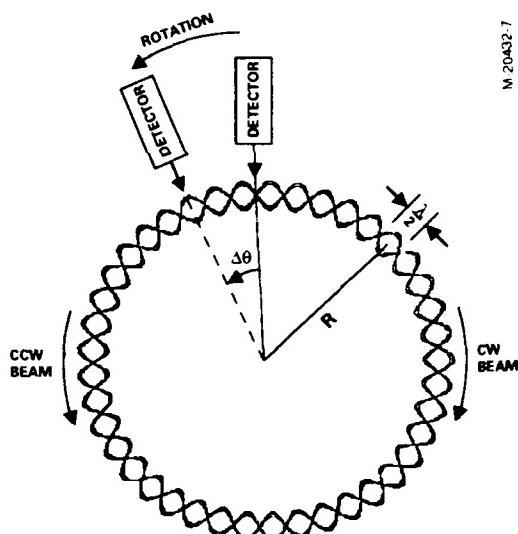
The RLG has undergone considerable development over the past two decades. The two-mode RLG is the most common optical inertial sensor while multioscillator gyros are showing much promise for the next generation of navigation systems. The following describes the principles of operation of these two optical gyroscopes. Many of the topics discussed apply to both forms of ring lasers.

4.1 Two-Mode Gyros

The first, and still most prevalent, ring laser gyros^{6,7} are two-mode resonators which support one cw and one ccw mode (the distinction between two-mode and multioscillator RLGs will become apparent later). A resonant cavity is formed with a block of glass possessing three or more sides and with very high-quality mirrors on each corner to close the cavity. Bores in the glass permit laser beams to propagate. The bores are also filled with a gas mixture (generally Helium-Neon) which serves as a gain medium. One or more cathodes and one or more anodes are used to support an electric discharge in the gas. This serves as the pump for the laser which then resides within the resonant cavity. The RLG is, therefore, an active resonator.

At least one of the mirror faces is partially transmissive to permit observation of the laser beams. Combining optics, typically in the form of a prism, are used to coherently recombine (heterodyne) the cw and ccw beams in order to permit the observation of the interference and hence the standing wave pattern (also referred to as interference fringes) created by the counter-propagating modes. Detectors are used to measure the intensity of the interference fringes.

As discussed in the previous section, the Sagnac effect renders the standing wave pattern stationary in inertial space. When the gyro is rotated, it is the body which rotates relative to the standing wave. Measurements of angular rotation are performed by observing the changes in the detected heterodyned outputs. The operation of the RLG is best illustrated in Figure 1 which depicts a fictitious circular RLG. In this figure, a virtual standing wave pattern is created by the two counter-propagating waves. When the gyro is rotated, the detector moves with respect to the interference pattern and dark and light areas will alternate. Each dark/light cycle represents one half wavelength of the laser beam along the circumference of the path. The number of dark/light transitions can therefore be geometrically related to the angle of rotation as indicated in Figure 1. By counting the number of fringes traversed, a measure of the total rotation angle is obtained. The scale factor of the instrument is based on purely geometrical relationships as given in Figure 1 for the "circular" RLG. A similar analysis can be made for any polygonal gyro.



- INTERFERENCE BETWEEN CW AND CCW BEAMS CREATES STANDING WAVE PATTERN
 - STANDING WAVE PATTERN STAYS FIXED IN INERTIAL SPACE
 - AS GYRO CASE ROTATES, DETECTOR MOVES AROUND RING AND COUNTS MINIMA
 - SCALE FACTOR CORRESPONDENCE
- | | | |
|--|--|---------------------|
| $\frac{1}{2}$ CIRCUMFERENTIAL DISPLACEMENT | $\rightarrow \theta = 2\pi \text{rad}$ | OPTICAL PHASE SHIFT |
| $\Delta\theta$ MECHANICAL ROTATION | $\rightarrow \theta = \frac{2\pi D}{\lambda} \Delta\theta$ | OPTICAL PHASE SHIFT |

Figure 1. Circular Ring Laser Gyro (RLG)

The scale factor remains geometrical, involving the ratio of enclosed area to path length according to the following expression:

$$N_{\text{cycles}} = \frac{4A}{\lambda L} \Delta\theta$$

where: A = Enclosed Area of Gyro
 L = Path Length of Gyro
 λ = Wavelength of Laser
 $\Delta\theta$ = Angle Increment

4.1.1 RLG Quality

The RLG makes use of the laser within the cavity to provide the signals which create the standing wave pattern and hence the measurement capability. As is well known, the laser is based on stimulated emission of photons. However, the gas medium which supplies the gain for the laser also occasionally emits photons which are unrelated to the laser signal. This is known as spontaneous emission and leads to noise and random walk in the RLG angle output. Spontaneous emission is described statistically through quantum mechanics and cannot be eliminated. However, in order to reduce its impact on gyro performance, it is important to maximize the active signal. This is accomplished by designing a gyro with high gain and low loss which results in a high "finesse" (analogous to Q in a resonator).

For the reason discussed above, it is exceedingly important to incorporate extremely high-quality mirrors into the RLG. Low loss is required to minimize the impact of spontaneous emission and hence reduce the "quantum limit" which is a measure of the best noise performance (and hence angle random walk) achievable with the gyro. For reasons to be discussed later, it is also essential to minimize the backscatter generated by the mirrors.

4.1.2 Technical Challenges in Two-Mode RLGs

The most severe problem encountered in the RLG is that of lock-in. At low rotation rates, it was observed that no fringe displacement was obtained. That is, the RLG was apparently insensitive to low rates. This is illustrated in the output versus input curve shown in Figure 2. The cause of the lock-in phenomenon is backscatter within the cavity, usually resulting from imperfections or particulates on the mirror surfaces. At low rates, the two counter-propagating beams in the resonator are very close in frequency because their optical pathlengths are nearly equal. Coupling of one beam into the other (which results from backscatter) causes the two modes to "lock" together thereby making the gyro insensitive to the actual rate. In the schematic representation of the RLG given in Figure 1, backscatter amounts to friction between the standing wave pattern and the cavity. When the gyro is rotated at low rates, the standing wave pattern "sticks" to the cavity instead of remaining fixed in inertial space. The detector therefore does not observe any shift in the interference fringes and the gyro does not register the rotation. At high rates, the "friction" is overcome and the gyro is capable of measuring rate.

4.1.3 Lock-in Circumvention

In a two-mode RLG, the method of overcoming lock-in consists of mechanically biasing the gyro to the point where lock-in becomes negligible. The conventional means for accomplishing this is the use of dither wherein a large amplitude sinusoidal excitation is applied to the gyro body. The angular rate corresponding to this motion typically is on the order of 100 deg/sec. The output of the gyro must then be compensated for the dither

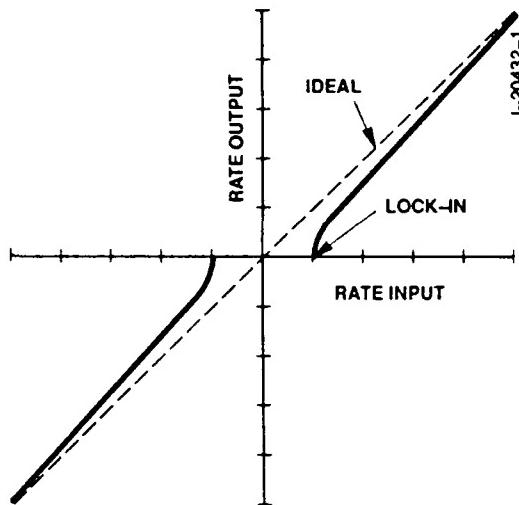


Figure 2. Two-Mode RLG Input/Output (No Dither)

motion so that the true rotation of the vehicle can be determined. There are many effective techniques for accomplishing this compensation.

One additional draw back of dither is increased random walk as well as mechanical problems described below. As the sinusoidal motion crosses through zero velocity, a small lock-in error usually occurs. Since the gyro reverses direction twice per dither cycle, a random accumulation of these errors occurs thereby leading to random walk. The random walk coefficient improves with the square root of dither rate but is usually the dominant source of random walk in a dithered RLG.

An alternative method (which avoids the random walk problem) of biasing the RLG makes use of large angle rotation using a turntable. An angular encoder provides the relative angle between the instrument cluster and its base. This technique is referred to as rate biasing. Rate biased systems with small path length RLGs have been delivered for missile applications and excellent accuracy demonstrated. Nonetheless, because of the mechanical complexity involved in rate biased systems, these remain much less prevalent.

4.1.4 Mechanical Design

As mentioned above, most RLG systems today employ dither as lock-in circumvention. Unfortunately, dither places serious constraints on the mechanical design of the system. In order to achieve the high dither rates required, high frequency (typically several hundred Hz), high Q mechanical flexures are used. Loss of dither energy to the outside world is detrimental in terms of achieving the appropriate motion. In addition, coupling of dither motion to mounting structures may have undesirable effects such as acoustic noise. Thus, hard-mounted dithered systems are generally not practical. A relatively low frequency suspension is usually used to isolate the sensor assembly from the outside world.

Another important consideration in the mechanical design is that coupling of dither torques between the three axes of the instrument block can lead to coning motion of the entire

assembly. This leads to apparent gyro bias errors in the navigation solution.

The mechanical design of a sensor assembly for dithered RLGs is therefore extremely critical and must be approached with considerable care. While state-of-the-art dither controllers tend to mitigate the problems due to non-ideal mechanical behavior, it must be highlighted that the mechanical aspect of dithered systems remains a very real and relevant challenge.

4.1.5 Cavity Length Control

The RLG operates as a resonant cavity. However, the gas mixture which supplies the pump for the laser exhibits a well-defined gain curve. Certain ranges of optical frequencies (or equivalent wavelengths) excite the stimulated emission and result in lasing action. In order to ensure proper operation of the gyro, it is essential that the cavity be tuned to the proper resonant frequency. That is, the length of the cavity must be an integral number of the proper wavelength. For the Helium-Neon gas mixture normally used, this wavelength is on the order of 630 nanometers. It is obviously impossible to mechanically establish a cavity whose length is accurate to a small fraction of this value by design! Furthermore, the thermal expansion or contraction of the glass used for the cavity will cause the length to vary as a function of temperature. The solution to this problem is the use of an active scheme in which the total laser intensity is monitored and the mirror positions adjusted to maximize this intensity. The mirror motion is created by piezo-electric transducers which causes minute displacements of the mirror faces. Nonetheless, because the excursion afforded by the mirror transducers is very small (typically only a few wavelengths), it is essential that the cavity be made of a low expansion glass so that the mirror travel is sufficient to compensate for expansion over the entire temperature range. If this is not the case, then "mode hops" must be performed wherein the path length of the gyro is quickly changed by one wavelength to establish a new control point. Unfortunately, a mode hop causes a loss or a corruption of data during the time it is occurring. Frequent mode hops or mode hops during high dynamics are therefore to be avoided.

4.1.6 Gain Medium

A gas mixture within the cavity supplies gain for the laser. The design of the RLG cavity should be such that gas flow within the gain bores remains balanced. A net flow around the cavity causes gyro bias and can be a dominant error source in any RLG. The gain medium is also a consideration in the instrument lifetime. In very small RLGs with unfavorable volume to surface ratios, Helium permeability can compromise lifetime.

4.1.7 RLG Error Models and Scaling Laws

The performance of a RLG depends very heavily on its size. The parameters used to describe gyro performance include the random walk coefficient, bias stability, resolution (also known as quantization), and scale factor stability. Because the sensitivity of the gyro is determined by the ratio of its area to its pathlength, it follows that all else being equal, the performance should improve at least linearly with pathlength. However, most of the parameters are actually much higher-order functions of length. The following scaling laws are provided as guidelines. Because of the interaction between many elements in the gyro design, these laws may not be perfectly accurate but they do illustrate the concept "a little pathlength goes a long way."

$$\text{Quantization : } \frac{1}{L}$$

$$\text{Bias Stability : } \frac{1}{L^3}$$

$$\text{Scale Factor Nonlinearity : } \frac{1}{L^{5/2}}$$

Random Walk :

$$\text{Spontaneous emission : } \frac{1}{L^2}$$

$$\text{Dither induced : } \frac{1}{L^3}$$

Reference 6 provides a more detailed discussion of error sources and mechanisms. It is noted that in the above expressions, the quality of the components which form the RLG are assumed to be independent of the pathlength. Thus, reducing the size of the RLG may not significantly reduce its cost since high quality components (mirrors in particular) are still required. In effect, quality becomes more critical as pathlength is reduced.

In addition to the scaling of the gyro parameters themselves, other performance considerations apply to the mechanical design of the gyro and its mounting structure. For example, random walk in a dithered gyro improves with the square root of the dither amplitude. However, as discussed earlier, mechanical cross-coupling between gyros can cause coning motion which appears as a navigation gyro bias. The coning error is proportional to the square of the dither amplitude. Thus, while it may be possible to improve random walk by raising dither amplitudes, the penalty incurred is a coning rate which increases rapidly with amplitude.

4.1.8 Vibration Sensitivity

While the RLG itself is insensitive to vibration, its mechanical configuration may lead to dynamic errors when it is used for navigation. For example, if the RLG flexure (which permits dither) exhibits any cross-axis compliance, then the set of three gyros in an inertial system no longer forms a rigid body and large navigation errors may result during vibration. This is a difficult problem because the flexure must be compliant (with a very high Q) about one axis while remaining extremely stiff about the other two perpendicular axes. A similar problem is encountered if the pickoff employed to measure the relative rotation of the RLG frame with respect to its case exhibits any sensitivity to acceleration.

4.1.9 State of the Two-Mode RLG

The two-mode RLG represents a mature rotation sensor, which in reasonable sizes, offers good bias stability, scale factor accuracy, low random walk, and extremely high dynamic range. RLGs with pathlengths on the order of 30 centimeters now form the basis for most modern, 1-nautical mile per hour navigators in the military and commercial markets. The two-mode RLG is evolving with reduced size devices now made possible by improved components and new error compensation techniques. Gyros with pathlengths of 15 to 20 centimeters now offer performance comparable to that of earlier 30-cm gyros.

4.2 Multioscillator Gyros

While the two-mode RLG was developed into an accurate device for measuring rotation, its need for mechanical biasing (i.e., dither or rate bias) has been problematic from a practical standpoint as well as an aesthetic one. Dither has now been perfected but early on it proved to be a major problem area in terms of gyro performance. Even at this time, dither imposes severe constraints on the mechanical designs of RLG systems and

causes increased noise, random walk, and physical motion of the sensor assembly. From an aesthetic standpoint, dither and rate bias are unattractive because they imply instruments with moving parts which are not truly strapped down to the vehicle. For these reasons, means of optically biasing RLGs have been attempted since the early days of RLG research. This has led to the class of RLGs known as multioscillators. These gyros are currently being delivered for use in commercial transports and military aircraft. They offer advantages which include very low noise, are truly "strapped down," and may be used as very wide bandwidth, very high resolution angle sensing devices. They have no moving parts and generate no acoustic noise. A description of their operation follows.

4.2.1 Operation of the Multioscillator RLG

As discussed earlier, the two-mode RLG is a resonant cavity which supports a standing wave pattern. Two-mode RLGs are planar by design so that only linearly polarized modes may be resonant in the cavity. Furthermore, in order to ensure stable operation, one of the two polarizations is suppressed. In a multioscillator gyro, both polarizations are used to form a left-hand circularly polarized (LCP) mode and a right-hand circularly polarized (RCP) mode which coexist within the cavity. In the multioscillator devices, these two modes are split apart in frequency. In effect, this corresponds to having two gyros active within the same resonator. Separation of the LCP and RCP modes is accomplished by making their optical paths different. This can be done with an optically active crystal in the beam path which rotates polarization states and consequently introduces a differential phase between the LCP and RCP waves⁸. A more attractive alternative makes use of image rotation⁹. When a set of mirrors creates a beam path which does not lie in a plane, an image rotates in space as it is reflected around the path. A net rotation may be obtained through a complete circuit. The image rotation also leads to a polarization rotation which splits the RCP mode from the LCP mode. The advantage of the image rotation mechanization is that it does not require an additional element to be present in the beam path thereby avoiding scatter and loss.

Once the LCP and RCP modes are split, they may be treated as two separate gyros each possessing clockwise and counter-clockwise beams. As such, lock-in may occur on each of the gyros thereby precluding rate measurements. In order to avoid this, a magneto-optical effect is used to bias the cw beam away from the ccw beam. This is generally accomplished with a doped glass element in the beam path which, when subjected to a magnetic field, causes a differential phase shift between the cw and ccw beams. Furthermore, the shift is in opposite directions for the LCP and RCP modes. The phenomenon responsible for the circular polarization dependent phase shift is known as the Faraday effect and the glass element responsible for its occurrence is known as a Faraday rotator. The frequency splitting in an out-of-plane multioscillator gyro is illustrated in Figure 3.

The multioscillator gyro measures rate in much the same way as two independent gyros (one LCP and one RCP). However, because of the built-in split between the cw and ccw modes, each gyro appears to be biased to a large rate (the Faraday bias) thereby avoiding lock-in. As illustrated in Figure 4, when true rotation rate is applied to the multioscillator, the rate sensed by one of the two gyros (LCP in Figure 4) increases while the rate sensed by the other gyro (RCP in Figure 4) decreases. By subtracting the outputs of the two gyros, it is possible to cancel the Faraday bias while *doubling* the true angular rate measurement.

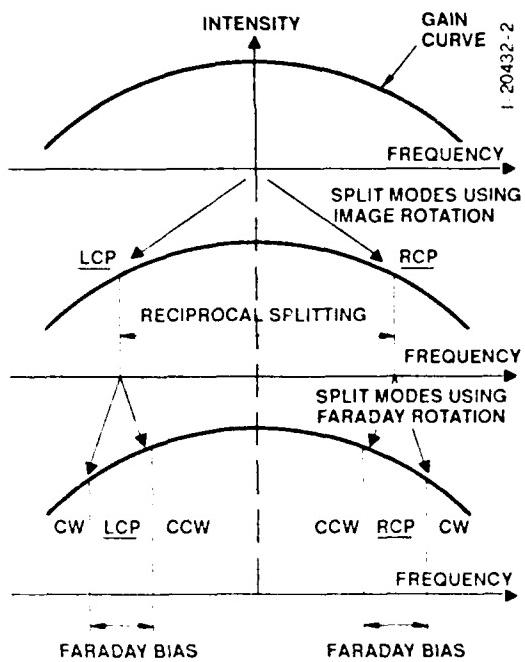


Figure 3. Mode Splitting in a Multioscillator RLG

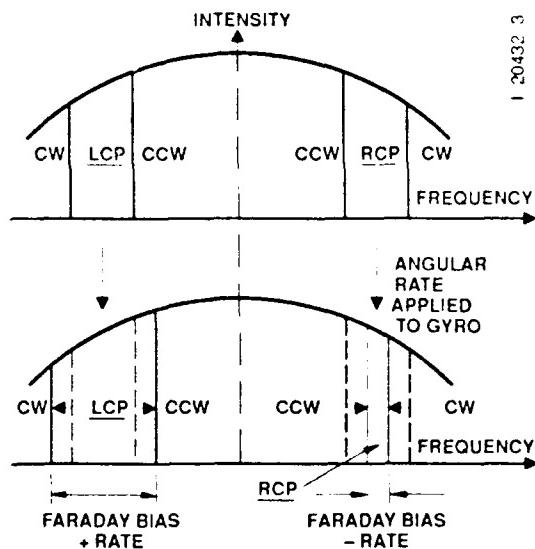


Figure 4. Effect of Rate on Multioscillator RLG

Thus, the rate sensitivity (scale factor) of the multioscillator is essentially double that of a conventional two-mode gyro of the same size but still determined by the geometrical ratio of projected area to pathlength.

4.2.2 Technical Challenges in Multioscillator Gyros

While the most significant problem associated with two-mode gyros (i.e., lock-in) is avoided in multioscillators, other problems arise because of the optical configuration. The Faraday split which is employed to separate the cw and ccw modes requires that a glass element be present in the beam path. Interaction between scatter sources on the surfaces of the Faraday rotator and the mirrors causes mode coupling which in turn, leads to gyro bias. However, by minimizing the amount of scatter through the use of state-of-the-art mirrors and rotator coatings and by employing designs which minimize sensitivity to scatter, the bias stability difficulties can be mitigated.

4.2.3 Advantages of the Multioscillator Gyro

While the strong point of the two-mode RLG has always been its bias stability, the multioscillator gyro presents many advantages. Elimination of mechanical biasing requirements for lock-in circumvention make this gyro exceptionally well suited for low noise, multifunction (flight control) applications. The elimination of dither leads to a random walk which approaches the quantum limit. The scale factor stability is exceptionally good due to the absence of residual scatter-induced lock-in effects present in dithered gyros. The doubling of the scale factor ensures that smaller instruments can be used, and the lack of dither-induced mechanical noise implies superior angle measurement and flight control capabilities. The mechanical designs are simplified due to the absence of high frequency, high-Q dither flexures. In addition, since the multioscillator is a true strapdown sensor, it may be used for accurate pointing and tracking applications as well as for navigation purposes.

4.2.4 Mechanical Design

As in the two-mode RLG, the gyro itself does not exhibit vibration or g sensitivity. A sound mechanical design is still required to avoid rectification of errors in the navigation frame. However, the restrictions incurred because of mechanical biasing in a two-mode RLG are eliminated in regards to the multioscillator. While cross-axis compliance must still be minimized, there is no simultaneous need for a built-in compliance about only

one axis. High frequency, dither-induced coning is also eliminated and the absence of dither implies that somewhat higher frequency isolators may be used.

4.2.5 Cavity Length Control

As in the case of the two-mode RLG, the multioscillator gyro cavity length must be controlled to ensure resonant operation within the gain curve provided by the gaseous medium. In addition, it is important that the LCP and RCP intensities be balanced to ensure common mode cancellation of errors when the output of the two gyros are subtracted. Cavity length control may be accomplished in several different ways. One method similar to that used in two-mode RLGs involves maximizing total intensity while others attempt to control the difference between the LCP and RCP intensities.

4.2.6 State of the Multioscillator RLG

While the multioscillator gyro has been under development for the past two decades, activity has been intensified in the past five years. Technology has progressed rapidly to the point where navigation grade instruments with pathlengths on the order of 20 cm have now been produced and delivered. Development continues on the multioscillator to further enhance bias stability and other key performance parameters. As in the case of the two-mode RLG, pathlength and component quality are essential to performance. The multioscillator's many benefits make it very attractive for present and future systems.

5. FIBER OPTIC GYRO

The second category of optical gyros discussed is that of Sagnac interferometers. The most viable form of the fiber optic gyro (FOG) falls within this category.

5.1 Principle of Operation

The interferometric fiber optic gyro (IFOG) is basically composed of a light source, a coupler, a fiber coil, and a detector. Figure 5 shows a generic configuration for an IFOG. Light is launched from a source and coupled through a fiber optic coil in both the cw and ccw directions. Because of the Sagnac effect,

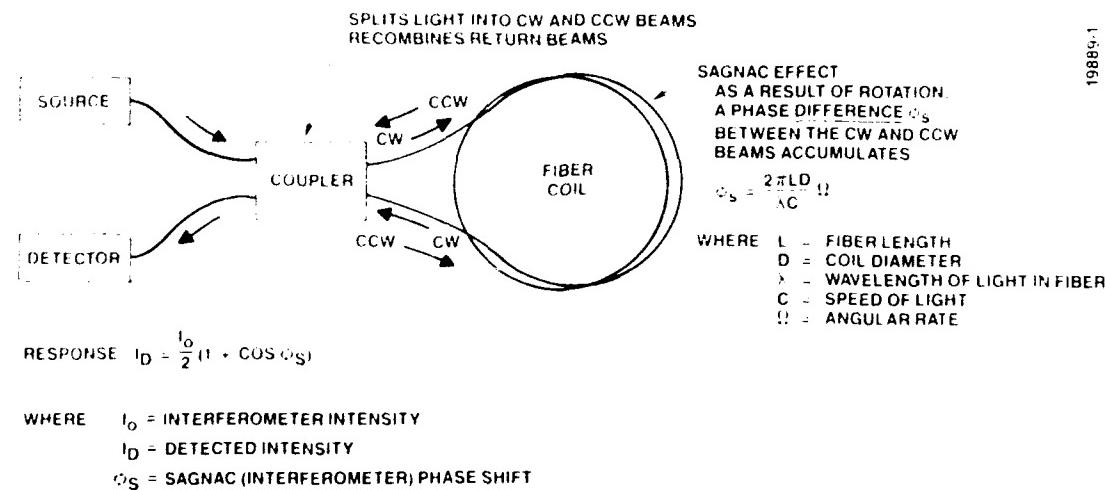


Figure 5. Interferometric Fiber Optic Gyro (IFOG)

the optical paths seen by the two beams differ in proportion to the rate applied to the gyro. Upon recombination, the two beams interfere and the intensity detected is a function of the phase difference and hence the rate of the gyro. It is shown, in Reference 3, that the phase difference is, in fact, proportional to the angle traversed during one transit time of light through the fiber coil and *not* the instantaneous rate as is commonly believed. Thus, the FOG's output characteristic is that of a rate integrating gyro with a very short memory as opposed to an RLG which is rate integrating over an indefinite period of time.

5.1.1 Fiber Optic Gyro Modulation

The interferometric architecture of the FOG implies that sensitivity is poor at low rates. This is illustrated in Figure 6 which

shows detected intensity as a function of rate input. In order to obtain good rate sensitivity, it is necessary to modulate the FOG to phase shifts which are odd multiples of $1/2\pi$. In early FOGs, this was accomplished by winding the fiber on a piezo-electric mandrel. An excitation was applied leading to a modulation of the light in the fiber. The development of integrated optics permits the replacement of the piezo-electric mandrel with an electro-optic modulator within the beam path as shown in Figure 7. Light passing through the modulator is phase shifted in proportion to the voltage applied. However, the modulator is not selective as to direction of beam travel. Differential phase shifts between the cw and cew beams are sustained for only one transit time of light through the coil. Thus, modulation must be applied every transit time. A detailed discussion of FOG modulation may be found in Reference 5.

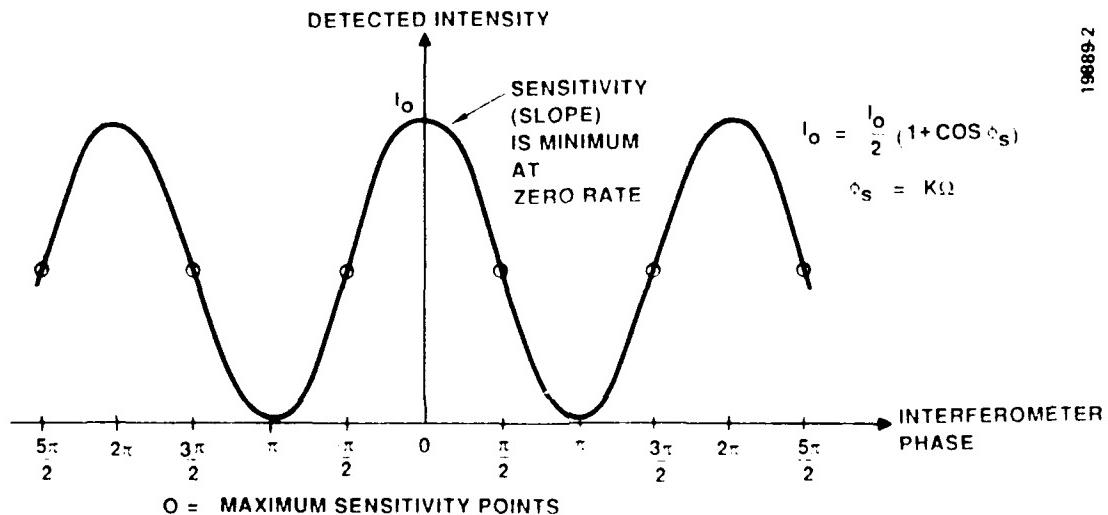
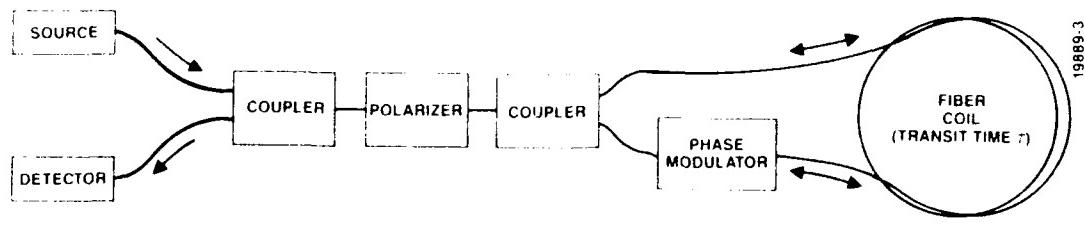


Figure 6. Detected Intensity vs Interferometer Phase



$$I_D(t) = \frac{I_0}{2} (1 + \cos(\phi_s + \phi(t) - \phi(t - \tau)))$$

WHERE
 I_D = DETECTED INTENSITY
 I_0 = INTERFEROMETER TOTAL INTENSITY
 ϕ_s = SAGNAC PHASE
 ϕ = PHASE MODULATOR PHASE
 τ = COIL TRANSIT TIME

Figure 7. IFOG with Phase Modulator

Modulation may take the form of a sinusoidal waveform. However, state-of-the-art devices often employ complex digital modulation to achieve maximum sensitivity and to avoid problems with distortion.

5.1.2 Closed Loop Fiber Optic Gyro Operation

Gyro modulation improves sensitivity at low rates. However, high rate nonlinearity, intensity variation, photodetector sensitivity, preamp gain, and background intensity all contaminate the open loop output of the FOG. For this reason, it is important for higher accuracy and greater dynamic range to operate the FOG in a closed loop fashion¹¹. This may be accomplished through the use of the same electro-optic device which supplies the phase modulation. Because a rate in the gyro appears as a interferometer phase shift, it may be nulled out by applying a phase rebalance (in addition to the modulation) with the modulator. However, the fact that the modulator can produce a differential phase shift between the cw and ccw beams *only* for a single transit time introduces a complication, because a given angular rate causes a *persistent* phase shift between the beams. Thus, in order to null out the rate induced shift with a phase modulator, it is necessary to increase the phase applied at least every transit time. Since available voltages are bounded, the phase cannot increase indefinitely. A periodic "reset" is required to maintain the voltage supplied to the phase modulator within prescribed limits. The magnitude of this reset must be exactly 2π of phase to ensure that the gyro is not perturbed. A block diagram of a typical closed loop FOG mechanization is given in Figure 8. As in the case of modulation, a digital implementation of the rate rebalance loop is attractive because it permits more precise control, tracking, and integration of the rate rebalance signal. A more detailed discussion of the closed loop FOG may be found in Reference 5.

5.1.3 Phase Modulator Scale Factor Tracking

The above discussion of the FOG describes the modulation and rebalance techniques. However, knowledge of the conversion between voltage applied to the electro-optic device and the actual phase shift generated is assumed. Modulation amplitudes equal to odd multiples of $1/2\pi$ and rebalance loop resets of 2π must be established. An error in the scale factor which relates phase to voltage may be observed by comparing the intensities measured in the course of $\pm 1/2\pi$ modulation against the intensities obtained during $\pm 1-1/2\pi$ modulation. A loop which measures the difference in intensities may be mechanized to control the voltage which corresponds to 2π .

5.2 Polarization Nonreciprocity

Many phase shifts in a FOG are reciprocal. This means that the shift experienced by the cw wave is the same as that experienced by the ccw wave and no differential phase is observable to contaminate the rate information. One error source which may be nonreciprocal is the coupling of different polarizations within the fiber/coupler circuit. This error often leads to an unmodelable uncertainty in the gyro bias and limits performance. Polarization nonreciprocity may be minimized through the use of high quality polarizers, short coherence length sources, and polarization maintaining fiber and/or depolarizers.

5.3 Vibration/Thermal Sensitivity

The index of refraction and the physical length of the FOG coil are affected by temperature and pressure. Differing conditions as a function of time as light propagates through the coil can lead to a rate error known as the Shupe effect. Typical conditions

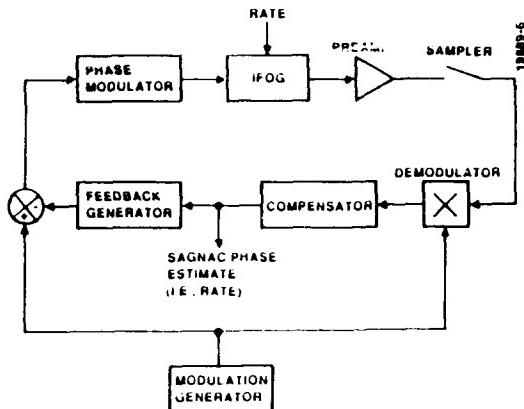


Figure 8. Closed-Loop IFOG Operation

which excite this phenomenon include a changing temperature-gradient or an inhomogeneous optical path subjected to varying temperature, or time varying stress resulting from a vibration input. Thermal Shupe effects lead to a gyro bias which is a function of temperature and gradient changes while mechanical Shupe converts ac translational vibration into an erroneous ac rate. Both of these effects may be reduced through coil winding methods which favor cancellation of Shupe errors. Thermal modeling techniques which account for temperature changes may be used to further improve performance while mechanical design can improve the mechanical Shupe effect.

5.4 Electronics

Short fiber length FOGs are made possible by very fast electronic components which are used to generate modulation, data processing, and loop closure. It is noted, however, that the electronics play an integral role in the performance of the FOG and that very careful design must be pursued. A digital approach to modulation, demodulation, and loop processing typically offers the most flexible and effective means of controlling FOGs while avoiding many of the instabilities associated with analog electronics.

5.5 Advantages of Fiber Optic Gyros

While the FOG is more demanding in terms of loop electronics as opposed to RLGs, it does present many advantages. The FOG requires no mechanical biasing, and is rugged enough to be operated in a hard mounted configuration. High-speed electronics permit very wide band operation with relatively short fiber lengths (on the order of 100-to-1000 meters). This leads to small size, weight, and cost without sacrificing performance. The FOG also provides extremely fine quantization thereby permitting its use not only as a rate integrating device but also as a very low noise rate sensor.

5.6 State of the Fiber Optic Gyro

IFOGs have been constructed and demonstrated for many applications. At the low performance end of the market, open loop gyros have been demonstrated for munitions. In the AHRS (Altitude and Heading Reference System) and small IMU arena, 1 deg/hr class systems have been flight tested, delivered, and are now beginning production. Navigation grade FOGs have also

been proven and form the basis for FOG navigators. It is anticipated that FOG technology will continue to evolve rapidly thereby leading to smaller, lighter, more cost effective inertial systems.

6. CONCLUSION

Optical gyroscopes form the basis for the majority of future INS'. The two-mode RLG is presently the most mature of the optical gyroscopes. Navigation systems which perform better than 1 nmil/hr are readily available both for military and commercial applications. Improvements to the component technology, the basic understanding of the physics, and the development of effective control mechanizations now permit the use of smaller RLGs.

One major drawback of the two-mode RLG is its need for mechanical biasing. The multioscillator RLG was developed to avoid this problem. Four-mode multioscillators are now capable of navigation grade accuracies and offer a number of advantages. This form of RLG has been demonstrated and deliveries have begun both for commercial and military programs.

Finally, the FOG with its small size and low cost has been demonstrated for AHRS class accuracy (1 deg/hour typical). FOG-based IMU's are being constructed for several different applications. Navigation grade gyros have also been successfully demonstrated and these will be integrated into the next generation of inertial navigation systems.

Each optical gyroscope discussed above offers unique performance capabilities to satisfy most of the requirements for inertial angle sensors. It is expected that they will remain dominant in the inertial navigation arena for many years to come.

7. REFERENCES

1. "Sagnac Effect," E.J. Post, **Reviews of Modern Physics**, Vol. 39, No. 2, April 1967.

2. "A Rate Integrating Fiber Optic Gyro," J. Mark, D. Tazartes, B. Fidric, A. Cordova, **NAVIGATION**, Vol. 38, No. 4, winter 1991-92.
3. "Fiber Optics Based IMU," D. Tazartes, SAE Technical Paper Series, Paper No. 901826, presented at Aerospace Technology Conference and Exposition, Long Beach, CA, October 1990.
4. "LFS-90 - A Modular System Design with Fiber Optic Gyros," E. Handrich, H.J. Büschelberger, M. Kemmler, M. Krings, **Symposium Gyro Technology**, Stuttgart, Germany, 1991.
5. "Optical Gyroscopes," D.Z. Anderson, **Scientific American**, Vol. 254, April 1986, pp. 86-91.
6. "Ring Lasers," J.R. Wilkinson, **Prog. Quantum Electronics**, Vol. 11, pp. 1-103, 1987.
7. "The Ring Laser Gyro," Chow, Gea-Banacloche, Pedrotti, Sanders, Schleich, Scully, **Rev. Modern Physics**, Vol. 57, No. 1, January 1985.
8. "Differential Laser Gyro Development," J. Bresman, H. Cook, D. Lysobey, Institute of Navigation Bicentennial National Aerospace Symposium.
9. "The Multioscillator Ring Laser Gyroscope," H. Stutz, T. Dorschner, M. Holz, I. Smith, **Laser Handbook**, Vol. 4, M. Stinch, M. Bass, Editors, Amsterdam, the Netherlands, 1985.
10. **Fiber-Optic Rotation Sensors and Related Technologies**, Springer-Verlag, 1982, S. Ezekiel and H. Ardity, Editors.
11. "Fiber-Optic Gyro Prototypes for High Dynamic Range Applications," H. Lefèvre, P. Martin, J. Morisse, P. Simonpietri, P. Vivenot, H. Ardity, **Symposium Gyro Technology**, Stuttgart, Germany, 1991.

Hemispherical Resonator Gyro: Principle, Design, and Performance

W.W. Stripling

Consultant

John R. Baskett

Engineer

Delco Systems Operations

6767 Hollister Avenue

Goleta CA 93117

USA

Abstract: The Hemispherical Resonator Gyro (HRG) is a new-technology rotation-sensing instrument based on the precession of a standing wave around a vibrating hemispherical form. This paper presents the principle of the HRG and describes the design and performance of some actual instruments.

Delco Systems Operations has developed a new-technology gyro, the Hemispherical Resonator Gyro (HRG), that is ideally suited to strapdown navigation systems. The HRG

tuned by the geometry of the resonator and the characteristics of the mode excited. From this it may be concluded that the HRG scale factor is largely independent both of rotation rate and of Young's Modulus for the resonator material. This latter implies low sensitivity of scale factor to temperature.

In applying Bryan's discovery to the development of a practical gyroscope, Delco selected the hemispherical shape for the resonator because it is inherently strong and readily manufacturable and can be supported on a shaft at its center, where flexing is minimal.

It is convenient to utilize the hemisphere's lowest-order resonant flexing mode, which consists of two pairs of antinodes. In the model discussed in this paper, the standing wave amplitude is less than 4 micrometers. Since, for a hemisphere, the higher-order resonant modes do not occur as integral harmonics, the lowest-order resonance can be driven without exciting higher-frequency harmonic vibrations.

Figure 1 shows how the alignment of the standing wave responds to rotation of the hemisphere. One antinode axis of the vibration pattern is initially aligned with a reference point. When the hemisphere is rotated, the vibration pattern precesses – it lags behind the rotation of the hemisphere by 30 percent. For instance, during a 90-degree counterclock-

The HRG is a Nearly Ideal Rotation Sensor

- "Solid-state" reliability
 - No rotating parts
 - Internal dissipation less than 10 nanowatts
- Insensitive to magnetic fields
- Nuclear ride-through capability
- Negligible acceleration sensitivity
 - Same performance on earth as in orbit
- High acceleration tolerance
 - Operates to 500 g
 - Survives to 3,000 g
- No temperature control required

offers great reliability and long life due to its inherent simplicity.

The Hemispherical Resonator Gyro is not a gyroscope in the classical sense, since it does not utilize a rotating mass, but only in the functional sense of being an inertial rotation sensor.

The rotation sensing property of a thin-walled resonator having the shape of a figure of revolution was discovered in 1890 by G.H. Bryan, a British physicist. His observations and their implications for the HRG are summarized below. Experimenting with wine glasses, Bryan demonstrated that the standing wave pattern of a ringing wine glass precesses around the bowl when the glass is rotated about its stem. He also showed, on theoretical grounds, that the ratio of the precession angle to the rotation angle is a constant deter-

Bryan's Discovery

- Wineglass experiment
 - Vibratory nodes precess when glass is rotated about stem
- 1890 paper (Proceedings of the Cambridge Philosophical Society)
 - Precession of nodes arises from Coriolis effect
 - Angular gain can be computed from geometry
- Implications
 - Wide-band, flat response
 - No rate threshold

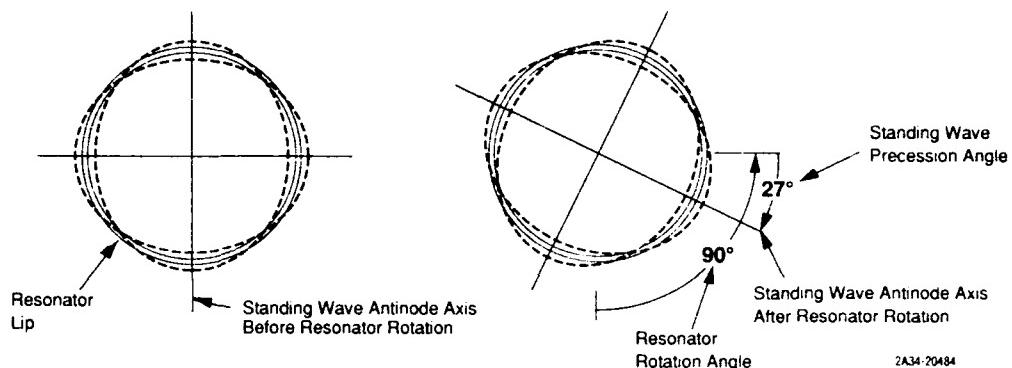


Figure 1. Precession of Standing Wave

wise rotation of the resonator, the antinode axis precesses clockwise relative to the resonator by 0.3 of the hemisphere's rotation angle, or 27 degrees. This scale factor of 0.3 is a geometric property of the resonator and holds constant for any rotation angle and any rotation rate.

The mechanical components of the HRG, shown in Figure 2, are the resonator, the forcer housing, and the pickoff housing. These three parts are all monolithic pieces made from fused quartz (silica) with metallized surfaces. The resonator is a hemispherical shell with a central shaft. The shaft is attached at one end to a forcer housing and at the other to a pickoff housing. In the assembled HRG, the outer surface of the resonator clears the electrode surface of the forcer housing by a few tenths of a millimeter, and the inner surface of the resonator clears the pickoff electrode surface by about the same amount. The resonator is driven and controlled electrostatically by the forcer electrodes, and the location and amplitude of its flexing pattern are sensed electrostatically by the pickoff electrodes. These elements of the HRG are enclosed in a vacuum.

Fused silica was selected as the structural material for the HRG primarily for its high "Q" and extremely low hysteresis. These material properties impart two important characteristics to the gyro: low damping, which is essential for high performance, and mechanical stability across vibration, shock, and thermal environments. Other desirable properties of fused quartz are its low thermal expansion and its great inherent strength when adequately surface-finished.

Three generations of developmental gyros, shown in Figure 3, have been constructed and tested. Each generation has achieved a substantial improvement in resonator damping time constant. The value of 1,800 seconds achieved with the third-generation gyros corresponds to a Q of about 12 million, thus validating the selection of fused quartz as the structural material. The tenfold reduction in damping losses obtained from the first to the third generation resulted from improvements in processes for surface preparation and metallization and in HRG geometry.

Growth in resonator time constant has played a major role in improving drift stability and reducing drift noise. Other

important evolutionary improvements have been gained in vacuum integrity, contamination control, and electronics.

To limit aerodynamic damping of the resonator, it is enclosed in a vacuum. Gases enter the internal volume from three sources: outgassing of absorbed gases from resonator and housing interior surfaces, leakage of atmospheric gases through imperfections in seals, and permeation of helium through the quartz housing. A gas getter is incorporated to maintain low internal pressure. Surface and particulate contamination are controlled during manufacture.

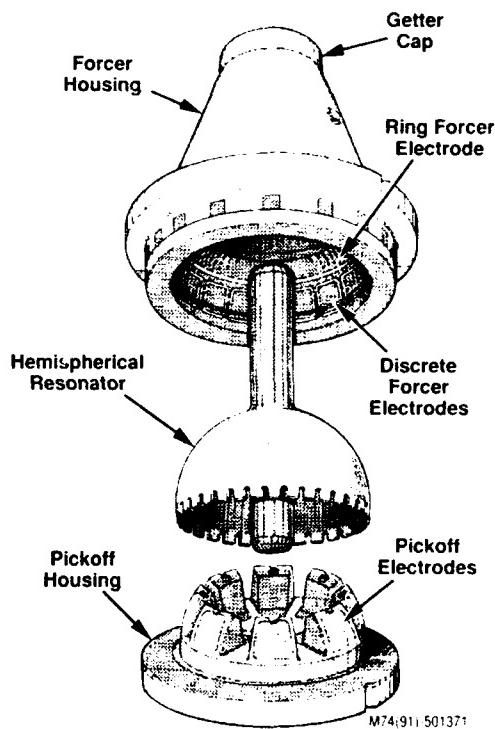


Figure 2. Mechanical Components of HRG

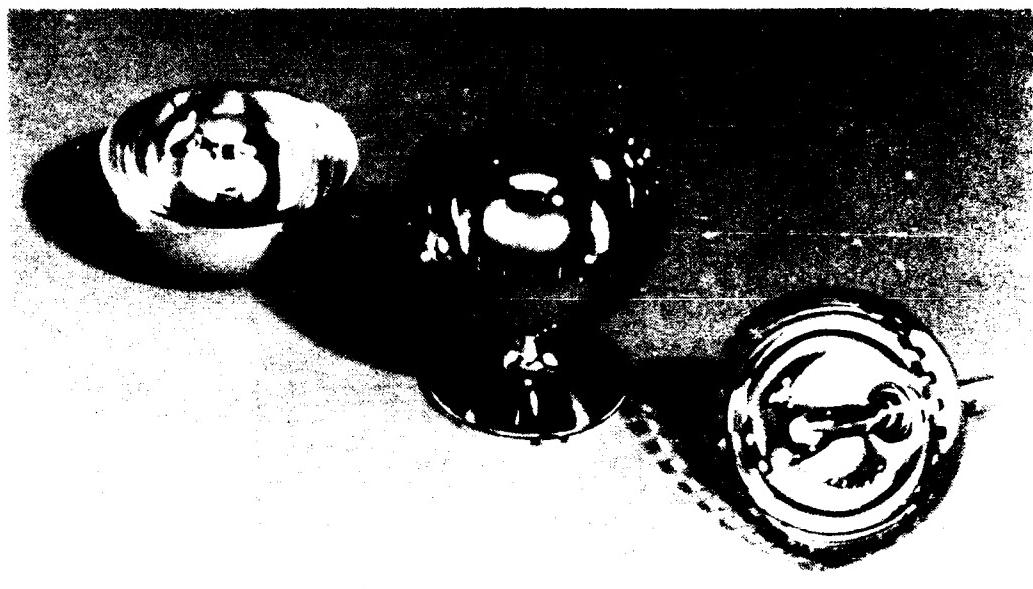


Figure 1. Three prototype of 4000- μ m-diameter resistive sensors for the particle detector.

the first development of the high-current resistive sensors for the particle detector. The first of these sensors was designed to be a permanent magnet. This type of sensor will be used to measure the particle momentum, and also for the readout electronics. In this work, the analytical experiments were developed specifically for the first design, and the first systematical study of the detector was undertaken. A single induction high-current circuit was obtained by using the electrical probes for the primary, open-circuited Hall-sensitivity transducers, which were not yet built.

1.2. The importance of the high-current resistive sensors

Development of a very high impedance buffer amplifier to isolate the resonator readout signals in order to obtain the high degree of linearity needed for readout readout efficiency, for the dynamic range.

Introduction of a trimmer process of the wave excitation loops to provide uniform behavior at all locations of the resonator standing wave pattern and during operation at high rotation rates.

The main electronic architecture is shown in Figure 4. The frame denotes the three principal control elements that maintain the resonator in the motion, as stable amplitude harmonic motion. These three elements and their role in maintaining the given performance are:

1. Amplitude regulation loop, which maintains a stable amplitude of the resonator standing wave pattern.

Frequency control loop, which maintains the frequency of the oscillation.

Density of the magnetic field, which maintains the track spacing without oscillations.

Resistive sensors for magnetic evaluation of the elements of the track and the readout signal at 800- μ m length.

3. Practical simplicity of low power consumption for readout and the simple forming of the standing wave all components in phase, the low-frequency.

The electronic hardware consists of a readout board, four current readout buffers, density control, magnetic control, digital processor, and power converter, as indicated in Figure 8.

The readout buffer amplifier is a low-performance device.¹ Four readout buffers are needed to form two balanced outputs with equal balance between channels; these four readout track over temperature to less than a few parts per million of full scale, and also provide linearity of the conversion.

The digital interface contributes to guide the readout A/D converter for principles of quadrature readout technique, which allows us to use the frame and the readout frequency counter, as well as to get the results of detection of the timer in a non-destructive way. The readout amplifiers and switches are used to feed the readout channel to the intermediate frequency, which can be correlated with the

¹ AD520, Analog Devices Inc.

Received 10 January 1990
Accepted 10 April 1990

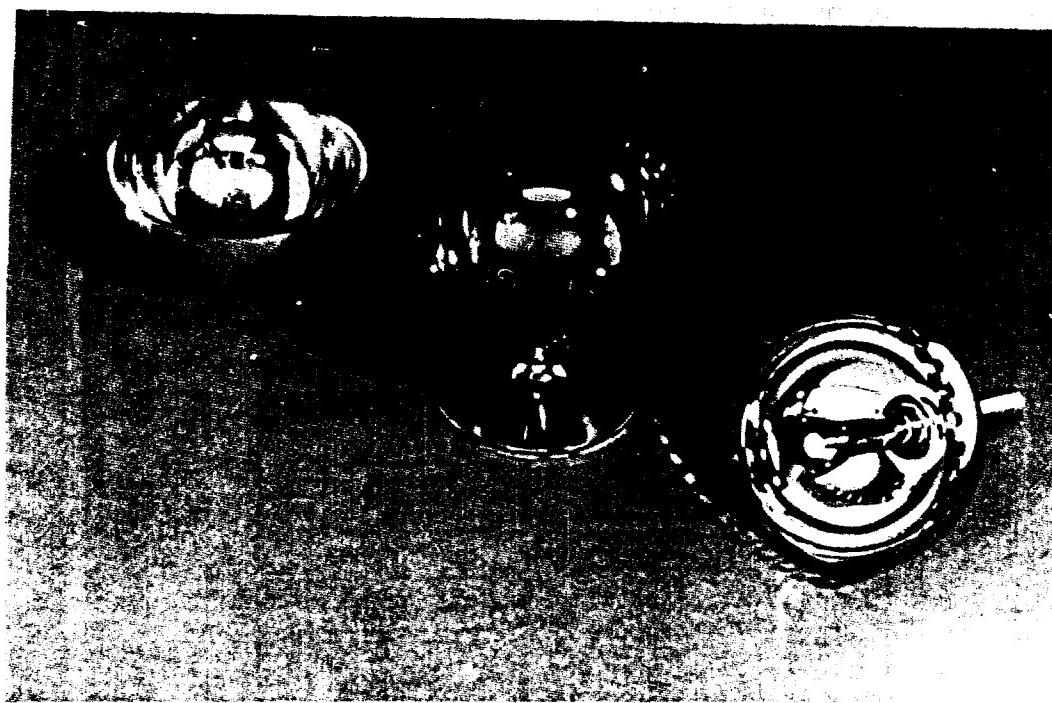


Figure 3. Three Generations of HRG's (left to right): First Generation, HRG 11; Second Generation, HRG 21-22; Third Generation, HRG 31-36

During the development of the HRG, performance has been electronics limited. Consequently, much of the evolutionary improvement has resulted from experimentally identifying a performance limitation, analytically modeling the phenomenon, and developing the needed electronic improvement. In this way, the electrical spring loops were developed to reduce drift by minimizing quadrature motion at the standing wave nodes. A major reduction in drift uncertainty was obtained by using the electrical springs also to provide open-loop compensation for drift sensitivity to nodal quadrature motion.

Two other important electronics improvements have been:

1. Development of a very-high-impedance buffer amplifier to isolate the resonator readout signals in order to obtain the high degree of linearity needed for accurate readout over an extremely large dynamic range.
2. Introduction of a digital microprocessor into the gyro control loops to provide uniform behavior at all locations of the resonator's standing wave pattern and during operation at high rotation rates.

The gyro electronics architecture is shown in Figure 4. The figure depicts the three principal gyro control elements that maintain the resonator flexing action as stable, simple harmonic motion. These three elements and their roles in maintaining the gyro's performance are:

1. Amplitude regulation loop, which maintains a stable amplitude of the resonator's standing wave pattern

2. Frequency control circuits, which develop timing signals used to
 - Demodulate the resonator's standing wave amplitude, spatial location, and phase
 - Replicate gyro flexing frequency, which is a direct measure of gyro temperature (frequency varies at 80 ppm/°C).
3. Electrical spring control loop, which is responsible for maintaining the simple harmonic motion of the standing wave (all components in phase at the same frequency).

The electronics hardware can be divided into four basic categories: readout buffers, digital interface electronics, digital processor, and power converters, as indicated in Figure 5.

The readout buffer amplifier is a key performance element. Four readout buffers are needed for each gyro. To maintain adequate balance between channels, these four must track over temperature to within a few parts per million of full scale, and also provide linearity of the same quality.

The digital interface electronics include the readout A/D converter for principal and quadrature mode detection from which error signals for timing and resonant frequency controls as well as flexing pattern orientation are derived; the timing signals generator, which includes the sampling and switching logic, and the D/A converter and driver circuits. Information from the readout signals is routed to the digital

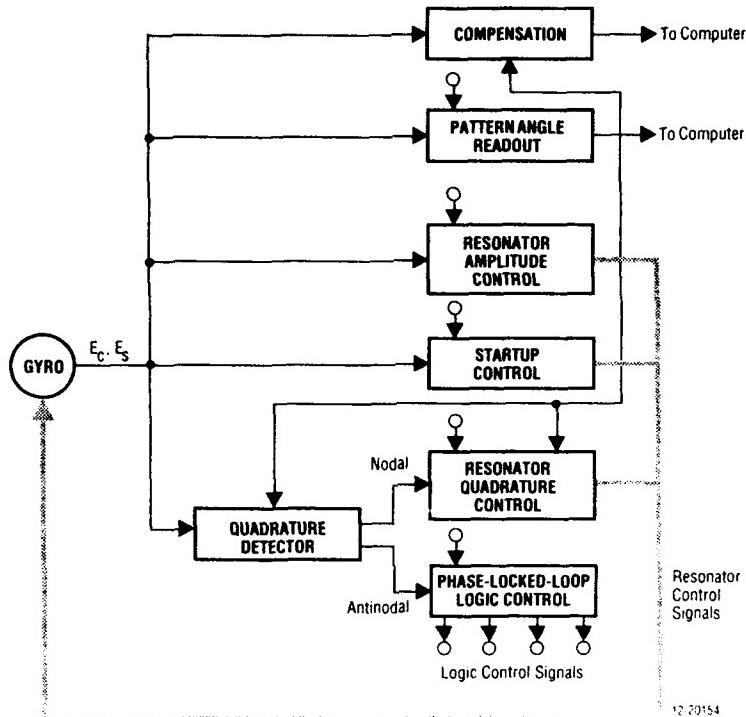


Figure 4. Electronics Architecture

processor through differential amplifiers that provide common mode rejection and are used to determine the amplitudes of the resonator's in-phase and quadrature flexing components and the flexing pattern precession angle.

An outstanding characteristic of the HRG is its very low level of angle random walk, which is the result of reduction in electronics-generated noise, especially that from power supplies and phase jitter in crossover detectors. Further reduction can be anticipated, because the theoretical limit is an order of magnitude below the operating performance of the current commercial instrument.

The HRG mechanical assembly is a simple, monolithic structure that has no wearout mechanism, operates at the sub-milliwatt power level, and is virtually immune to environmental stresses. Accordingly, HRG reliability is determined almost entirely by its electronics. The electronics are composed of low-power elements, including microprocessors and semi-custom interface circuit elements. With a maximum required voltage of less than 300 volts, these electronics are also expected to provide very high reliability.

The dominant gyro life consideration is a gradual increase in internal pressure, resulting in increased demands upon the gyro electronics. All electronics are designed to operate with a gyro time constant as low as 100 seconds. The gyro performs to specifications over this entire range of time

constant values. The operation of the gas getter assures a minimum 20-year life before performance degrades beyond requirements.

The principal HRG errors arise as a result of interactions between the gyro's mechanical assembly and its electronics. The mechanical assembly is characterized by a nominal axisymmetric configuration having small physical deviations such as lack of resonator/housing concentricity or circumferential variations in resonator thickness; the electronics compensate for such physical deviations. The electronics themselves are characterized by nominal behavior with small slowly changing deviations as well as small rapid fluctuations. Performance errors arise primarily from the small electronics deviations interacting with the nominal characteristics or deviations of the mechanical assembly.

Extensive vibration testing has confirmed that vibration-sensitive drift occurs only for vibration frequencies near the bending and flexing resonances, viz., the resonator flexing frequency (approximately 2,500 Hz) and the resonator/stem bending frequency (approximately 4,000 Hz). It also showed the need for a stiffer, noncantilevered resonator support for higher-performance gyros, which led to the double-ended shaft resonator configuration.

Family characteristics for resonators 58mm in diameter, 30μm in diameter, and 15mm diameter are presented in Figure 6. With the exception of calculated acceleration

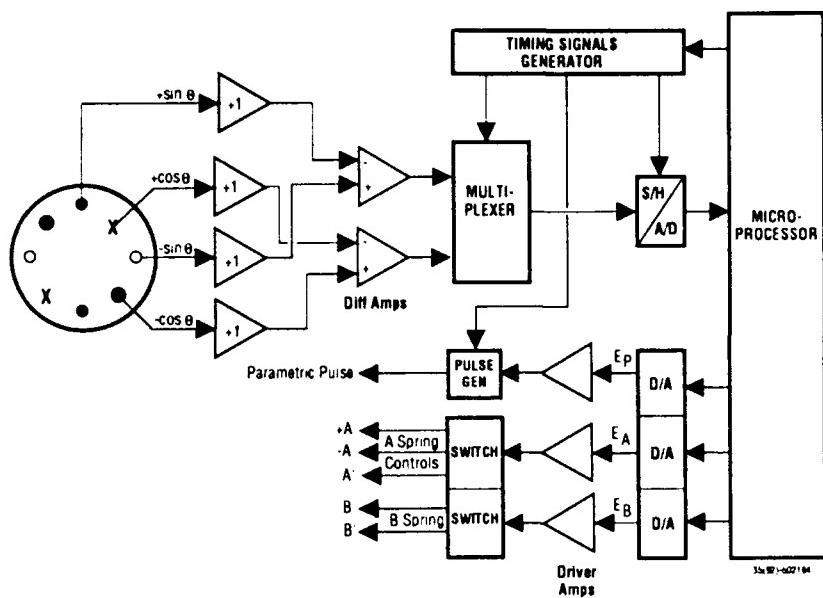


Figure 5. Electronics Hardware Design

CHARACTERISTICS	VALUES (1-SIGMA)		
	Actual	Projected	
	Full-Scale HRG (HRG 158)	Half-Scale HRG (HRG 130)	Quarter-Scale HRG (HRG 115)
Resonator			
• Flexing frequency (Hz)	2,500	5,000	10,000
• Bending frequency (Hz)	3,900	7,000	15,000
• Damping time(s)	1,800	1,000	110
Gyro			
• Max. diameter (mm)	91	47	25
• Max. height (mm)	101	63	35
• Weight (gm) (Includes gyro-mounted electronics)	544	82	50
Performance (1-sigma)			
• Drift uncertainty			
- Angle-dependent (deg/h)	0.002	0.01	0.7
- Bias (deg/h)	0.0001	0.001	0.06
• Angle random walk (deg/sqrt h)	0.00004	0.0001	0.002
• Scale factor uncertainty (ppm)	0.02	0.02	0.02
• Acceleration sensitivity			
- (deg/h/g)	2×10^{-7}	1×10^{-6}	8×10^{-6}
- (deg/h/g ²)	4×10^{-9}	1×10^{-8}	8×10^{-7}
• Readout accuracy (arc-seconds)	0.4	1.0	10.0

M-451-802366E

Figure 6. HRG "Family" Characteristics

Benefits of the HRG

- Intrinsically simple, low parts count
- No wearout mechanism
- No "moving parts" (resonator oscillation is less than 4 micrometers)
- No temperature control required
- No high voltage
- Potential failure mechanisms are minimized
 - Metallization failure – Processes monitored, parts tested
 - Contamination – Assembled in cleanroom
 - Loss of vacuum – 20-year (minimum) design goal; vacuum effects continuously monitored and compensated for
 - Drift coefficient stability – Gyro inherently stable; on-pad/on-aircraft calibration provided for electronically
- Outputs whole angle of rotation
- Resonator continues to measure rotation during power interruptions

sensitivity, values for the 58mm diameter resonator are from test data. Theory was used to project performance into half and quarter size instruments; gyros with 30mm resonators are currently being fabricated for test.

In summary, the HRG represents a revolutionary advance in inertial technology and provides the combination of characteristics needed in the next generation of inertial reference systems.

It is a strapdown gyro that meets the multifunction needs of advanced integrated avionics architectures, providing navigation accuracy beyond that presently available and also providing a wide-band, low-noise attitude reference for tracking, pointing, and flight control.

It exhibits very low error sensitivity to military operating environments, including highly dynamic environments (acceleration, vibration, shock, and angular velocity) and temperature extremes. It also is able to continue accurately sensing rotation during disruptions such as those caused by nuclear radiation.

It can respond rapidly from a completely dormant state, coming to stable operation within a few seconds. It has no significant thermal drift transient and requires no temperature control. It operates with very low random noise, a requirement for rapid, accurate gyrocompassing.

It has inherent characteristics making for low cost, long life, and high reliability; it is a solid state device having no wearout mechanism; the parts are fully interchangeable and suitable for automated assembly; and it operates at low power and low voltage.

The basic technology offers size/performance/cost tradeoffs that make it adaptable to a wide range of applications, including very accurate navigation for aircraft and strategic missiles, precision pointing and tracking for directed-energy weapons, small guidance units for ballistic missile interceptors, tactical missiles, and reentry vehicles, and ultra-small attitude references for hypervelocity projectiles and missiles.

Fiber Optic Gyros and Systems

H.-J. Büschelberger, E. Handrich

LITEF GmbH
Lörracher Str.18
7800 Freiburg
Germany

Summary

Fiber Optic Gyroscopes have emerged from the engineering laboratories and come into production. Conventional gyros will be replaced by the fiber optical gyros in most applications in the near future. LITEF has developed a family of gyros and systems based on these gyros. The introduction of the new technology brings many advantages with respect to function, performance, size and weight.

LITEF's Gyro Family

The LITEF fiber optic gyro (FOG) development program which started in 1983 has resulted today in a family of gyros covering a wide range of accuracy requirements. Fiber optic gyros are the gyro technology of the future. They have clear advantages over conventional mechanical gyros:

- no moving parts
- no acceleration sensitivity
- robustness
- high shock immunity
- high reliability
- low power consumption
- low manufacturing cost

These advantages make it quite clear that the fiber optic gyro will rapidly replace the mechanical gyro in most applications in the near future. LITEF has recognized the importance of this technology and has developed a family of gyros which meet market requirements. This FOG family comprises three variants, namely the K-2010, K-2020 and K-2030.

The highest performance gyro, the AHRS quality model K-2010, has been built in a prototype series over the past few years. Gyros of this model type were integrated in the FOG Attitude and Heading Reference System (AHRS), which was a modified version of LITEF's standard LTR-81 AHRS for commercial airlines. In May 1989 this FOG-AHRS was successfully flight tested at the DLR in Braunschweig (Germany). The results were comfortably within the requirements of the ARINC-705 specification in all points and on

all flights. In the same year, the LITEF K-2010 FOG instrument was tested in a comparative test-program sponsored by the German government and performed by the independent IABG test facility at Munich. In this test-program the LITEF fiber optic gyro emerged as the clear winner.^{1), 2)}

The other end of the performance spectrum is covered by the K-2030 which resulted from a development program whose aim was the achievement of high component packing density. The K-2030 met its project goals both in physical size and performance. This instrument will find use in low accuracy space-critical applications.

The K-2020 lies in the middle of the performance range where it meets the requirements set by an augmented AHRS system. It is smaller than the K-2010 in order to satisfy the requirements for low volume.

Based on its experience with the K-2030, LITEF started development in February 1990 of a three axis fiber optic gyro Inertial Sensor Unit for application in a flight control system. The restriction on size of the unit required a considerable miniaturization effort. This product is designated LFS-90. This system is described in more detail in the chapter "LFS-90 System Design".

All LITEF fiber optic gyro systems operate with a closed loop signal processing technique. This offers various important advantages over open loop operation:

- high scalefactor linearity
- high bandwidth
- unlimited input rate

The signal processing technique is covered by a number of special LITEF patents.

The LFS-90 unit was the basis for a family of FOG systems:

- LFS-92 3 axis rate gyro and accelerometer unit
- LFS-93 single axis rate gyro
- LCR-92/ μ AHRS attitude and heading reference system with level sensors
- LCR-93/ μ AHRS attitude and heading reference system with micromechanic accelerometers

A description of the functions and applications of these systems is given below.^{3), 4), 5)}

LFS-90 System Design

The LFS-90 is equipped with three fiber optic gyroscopes orthogonally oriented to each other on a common mechanical structure. The optical architecture is a triad configuration with one common lightsource for the three gyro sensor-axes, a coupler array, three sensor modules and three photodetectors.

The LFS-90 consists of five modules: three Sensor-modules, the Opto-module and the Electronics-module.

The Sensor module is connected to the Opto module via a fiber optic link. The Electronics module contains all the components required for three gyro axes.

Sensor Modules

A Sensor module is the sensing element of the gyro. It comprises in essence the fiber optic coil and a Multifunction Integrated Optics Chip (MIOC). This MIOC realizes the functions of the polarizer, the beamsplitter and the electrooptic phase shifter and is connected at one end to the fiber coil and at the other to the fiber link for connection to the Opto module.

Opto-Module

The Opto-Module functions as the interface between the optical and the electronic components of the gyro. It contains the semiconductor light source and the photodetector with its preamplifier. The current supply for the lightsource is placed on the circuit board of the Opto-Module and the read out couplers which are necessary to guide the returning light to the photodetectors are located in the Opto-module. A preamplifier converts the photo current to a voltage which is processed in the Electronics Module.

Electronics Module

The following functions are implemented in the Electronic Module:

- Modulation signal generation
- Closed loop Sagnac phase compensation
- Data filtering
- Interface control
- Self-test

The Electronics module is divided into the following three main sections:

- * Sensor Control Section
- * Processor Section
- * Interface Section

These sections contain the functions as described below:

* Sensor Control Section

The Gyro Loop Control is laid out to operate with three fiber optic gyros. The incoming signals from the Opto Module are A/D converted and processed in a digital ASIC which realizes all control loop algorithms. The output signals of the ASIC are D/A converted and applied to the modulator of the MIOC in the Sensor Module. The ASIC performs the main control loop directly in hardware to provide the fastest possible rebalance.

* Processor Section

A digital signal processor performs the data processing algorithms and supports the required interface functions. The tasks performed are in particular :

- Acquisition and processing of the measured sensor data from the ASIC. Compensation of these values with the calibration data resident in an EEPROM, and transfer of the required information via the dedicated interface.
- Performing power-up, background and initiated Built-In-Test functions in order to guarantee correct operation after power up and to detect failures during operation.

The operational program is stored in EPROM memory and the required calibration data is held resident in EEPROM memory.

* Interface Section

The Interface Section provides the required link between the Inertial Sensor Unit and flight control computer. This is realized by implementing two UART interface section into the ASIC. The interface is controlled by the digital signal processor.

Performance Data of the LFS-90

Total Drift Error	200 °/h	(3σ)
Drift Stability	24 °/h	(1σ)
Total Scalefactor Error	1 %	(3σ)
Scalefactor Repeatability	500 ppm	(1σ)
Data rate	60 +/- 5 Hz	
Power Consumption	12 W	
Size	160 x 100 x 35 mm	
Weight	< 700 gr	
Output Format		
electrical	RS 485	serial
addressing		hardware select line
sensor information	3 x rotation rates	
	time tag	
	status word	

LFS-92 System Design

The LFS-92 Inertial Sensor Unit is an extension and improvement of the LFS-90 ISU. This unit contains in addition to the fiber optic gyros, three micromechanical accelerometers. The data of the accelerometers are also processed by the instrument digital signal processor to provide velocity increments. The measurement data are error corrected and processed synchronously with the gyro data. The gyro and accelerometer data are both available at the system interface in the same data package:

- gyro data three 16 bit words
- accel data three 16 bit words
- status word one 16 bit word

The interface is a digital RS-485 serial link with a data rate of 50 Hz. This system has a built-in power supply to run the unit with 28 V DC.

Performance data of the fiber optic gyros are as follows:

Gyro Drift Error	36 °/h	(3σ)
Gyro Drift Stability	4 °/h	(1σ)
Total Scalefactor Error	0,5%	(3σ)
Input Range	2000 °/s	

Included in this system are three accelerometers of Litef's model B-290. This is a servo loop pendulous sensor with a differential capacitance pickoff. The servo loop operates with electrostatic caging of the pendulum. The whole sensor is made from silicon and is processed with similar technologies to those used in semiconductor manufacture. This makes automated batch fabrication of miniaturized sensors possible. Approximately 200 accelerometers are placed on a silicon wafer. The accelerometer is mounted together with its associated electronics in hybrid housing. The output of this hybrid is a digital signal.

The performance data of the B-290 accelerometer are as follows:

Size	25 mm x 25 mm x 6 mm
Measurement Range	+/- 10 g
Scalefactor	
Repeatability	< 1000 ppm (1σ)
Stability	< 200 ppm (1 σ) short term
Bias	
Repeatability	< 2000 μg (1σ)
Stability	< 500 μg (1σ) short term
Axis Misalignment	
uncompensated	20 mrad
compensated	200 μg
Power Supply	+/-15 V, + 5V

The physical characteristics of the LFS-92 unit are:

Size	52 x 100 x 160 mm
Weight	1,1 kg
Power Consumption	20 W

LFS-93 System Design

The LFS-93 fiber optic rate gyro unit is a single axis derivative from the three axis LFS-90 Inertial Measurement Unit. It comprises one fiber coil and the associated optics and opto-electronics components. Like all other LITEF fiber optic gyros this instrument also operates with the closed loop electronics. The performance data of the gyro and the electrical interface are identical to those of the LFS-90 unit. Differences in other parameters are as follows:

- power consumption 6 W
- size 107 x 88 x 34 mm
- weight 600 g
- no enclosing housing

The sensing axis is aligned parallel to the surface normal of the 107 x 88 mm area. The unit is specially designed to be integrated into a customer system. This is the reason why an individual housing for the gyro unit has been omitted.

Optionally the gyro axis can be oriented parallel to the long axis of the assembly. The dimensions and the FOG coil form are different in this case.

The switch-on time of the LFS-93 is less than 70 msec. The extensive self test routine of the LFS-90 unit after system start has been deactivated here to achieve fast reaction time. The background self test routine is continuously running. It monitors the power supply, the sensor temperature, the control loop and the output data.

The software protocol has additionally to the LFS-90 unit a simplified communication mode. The LFS-90 unit can decode 8 different commands and respond with the required information. The data exchange is described in a Interface Control Document (ICD). The LFS-93 gyro is usually operated in the data mode which sends gyro data with a fixed datarate when it is activated. It is LITEF's standard that the systematic error sources in the gyro output data are error corrected by means of the internal signal processor. Consequently temperature sensing and associated data processing for temperature modelling in the customers computer is not necessary.

LCR-92 / μ AHRS System Design

The μ AHRS measures the rotation rates of the vehicle and calculates its attitude angles relative to the horizon and heading relative to the north direction of the navigation coordinate system by quaternion integration. The system is equipped with bubble level sensors as a vertical reference. An external flux valve magnetic sensor unit is required for pre-flight alignment and heading augmentation. The body rates and body angles are delivered to the aircraft avionic system.

The μ AHRS is designed to meet the following standards:

- TSO C4c Bank And Pitch Instruments
- TSO C5e Direction Instrument, Non Magnetic
(Gyroskopically Stabilized)
- TSO C6d Direction Instrument, Magnetic
(Gyroskopically Stabilized)

The qualification is in accordance with RTCA DO-160 C including HIRF and lightning.

This system is specially designed for applications where low weight and small size are essential requirements.

size	276 mm x 105 mm x 130 mm
volume	3.2 liter
weight	2.5 kg
power	25 watt from 28 V DC

The performance data of the μ AHRS are listed blow:

attitude angle	0.5°	max	static
	2.0°	2σ	dynamic
heading			
augmented	2.0°	2σ	dynamic
DG mode	24°/h	max	incl. earth rate
body rates	0.1°/s	2σ	max 1%

These numbers apply to the digital outputs. The system is equipped with an ARINC 429 databus and optionally with a synchro interface. On the ARINC 429 databus the following labeled output data are available:

- Magnetic Heading
- Pitch and Roll Angle
- Body Pitch and Roll Rate
- Body Yaw Rate or Pseudo Turn Rate
- Platform Heading
- Discretes

As an input to the system is necessary:

- Flux Valve Input
- Discretes

The allowed input ranges for body angles and rates are:

pitch	+/- 90°
roll	+/- 180°
heading	0 - 360°
body rates	+/- 2000°/s

Alignment

Pre-flight alignment is performed automatically after switch on of the system. Valid data are available 30 seconds after switch on. The alignment can be executed during flight as well as on the ground. During an in-flight alignment procedure the aircraft must fly straight and level.

Magnetic Field Sensor

A standard flux valve magnetic sensor unit (MSU) is used for augmentation of the μ AHRS. The excitation of such a sensor can be delivered from the system. The appropriate output is available. For flights in regions with strong magnetic anomalies, the MSU input can be removed by a compass controller and the system then operates in the DG mode. The input for the compass controller is available in the system.

Calibration of the MSU is performed automatically by the μ AHRS. For this procedure the pilot has to fly a special maneuver which takes 4 minutes maximum. The calibration data which are calculated during the maneuver are stored in a special MSU calibration memory. This CALPROM is an external module which is plugged into a connector on the frontpanel of the unit. The module is linked to the tray and in case of removal of the μ AHRS from the aircraft this calibration memory can easily be transferred to another system. A new calibration routine of the MSU is not necessary after change of AHRS boxes.

LCR-93/ μ AHRS System Design

The function of this system is the provision of pitch, roll and stabilized magnetic heading for use by aircraft displays and other avionic systems. Additionally to the LCR-92 this system provides linear acceleration. The system comprises three axes of micromechanical accelerometers which replace the two axes level sensors.

The system has inputs from the magnetic sensor unit (MSU), the air data computer (ADC) and from two VOR/DME stations. By processing these data, the LCR-93 delivers the aircraft ground speed, wind speed, wind direction, flight path angle, drift angle, track angle and NS-, EW-velocity.

The system has three modes of operation:

- alignment
- normal mode
- DG mode
- basic mode
- compensation mode

The system automatically performs the alignment routine after switch-on in two phases. Phase one fulfills selftest functions and coarse alignment. After 30 seconds valid body rates, accelerations, attitude and heading data are available. Phase two of the alignment with the specified accuracies is completed after further 30 seconds.

The normal mode requires true airspeed data and magnetic heading input for augmentation.

The directional gyro mode (DG mode) is activated whenever the discrete input to the system is set to the DG position. The magnetic heading augmentation is deactivated in this case. The heading output of the system is subject to the gyro's drift error. The switch is also activated by software to disable the input when erroneous magnetic heading data are detected.

If true airspeed data are not available the system will automatically revert to the basic mode. The basic mode affects the accuracy of the pitch and roll angle output.

As described for the LCR-92 unit the magnetic sensor unit (MSU) has to be calibrated in the aircraft. The calibration routine and calculation of the compensation coefficients is performed in this operation mode. The coefficients are stored in the MSU-CALPROM module.

Fiber Optic System Applications

The family of Inertial Measurement Units as described above have applications in the following areas:

LFS-90

The LFS-90 is a three axis rate gyro unit for flight control or damping applications. It has been developed for an RPV program in Germany. The performance parameters and the small size fulfill the requirements of drones, missiles and comparable applications. The size and the weight of the unit make it one of the most attractive devices on the world market.

LFS-92

This IMU is similar to the LFS-90 but has in addition three micromechanical accelerometers. The fiber optic gyros have a lower drift error compared to the LFS-90 gyros. The application of this system is the high accuracy flight control where acceleration data are required. The small size, low power consumption and low weight are attractive parameters which fulfill all typical requirements of RPV's and other aircraft.

LFS-93

This single axis fiber optic gyro unit is the ideal sensor for most rate gyro applications. Its closed loop operation, short switch-on time and digital interface makes it a powerful instrument for a large variety of users. Today this instrument is installed in a navigation system of autonomous guided vehicles for factory transportation equipment.

LCR-92/ μ AHRS and LCR-93/ μ AHRS

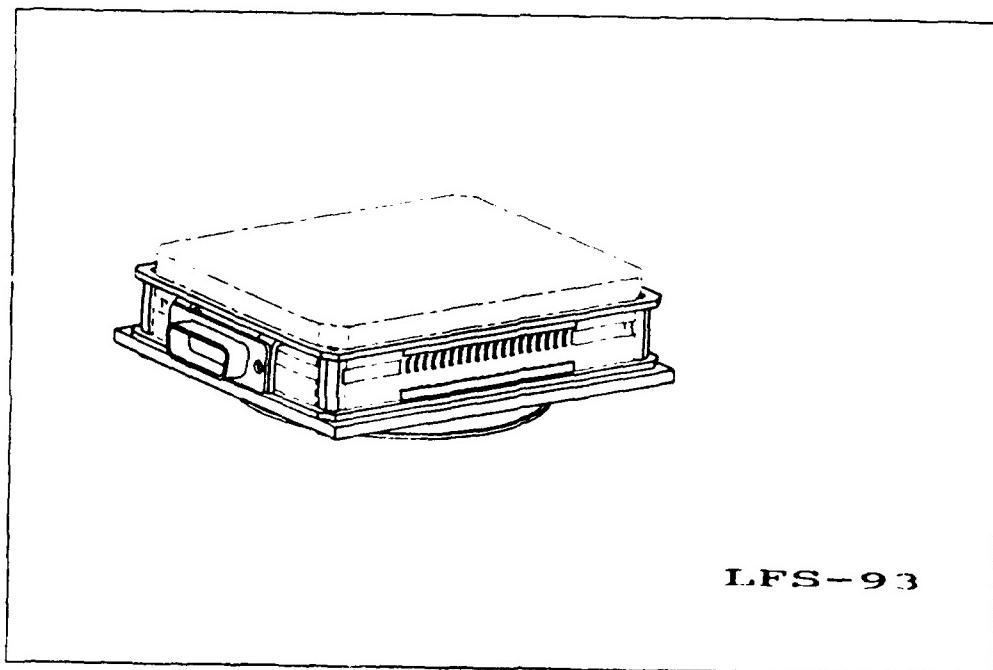
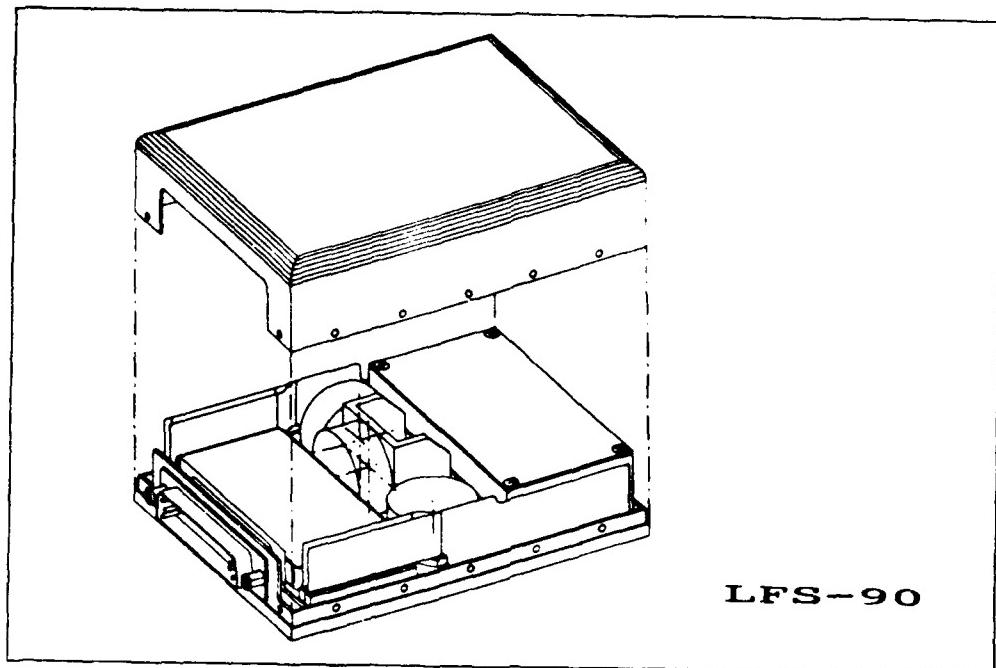
The μ AHRS is a low cost, low weight and small size Fiber Optic Inertial Measurement System which delivers attitude and heading information for aircraft avionic systems and can replace conventional VG/DG installations in aircraft. These systems are designed primarily for use in:

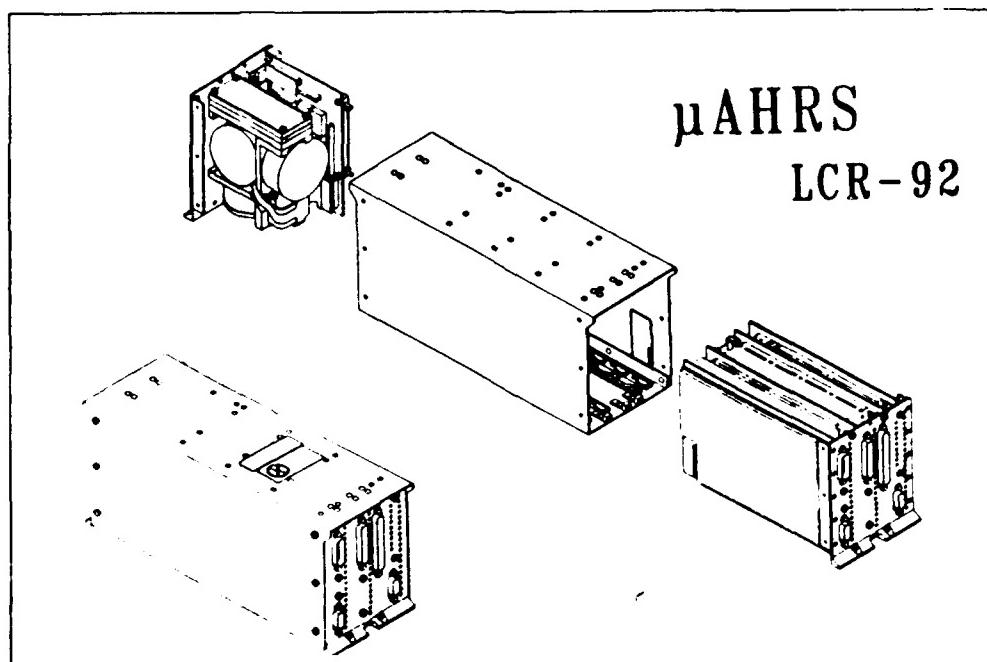
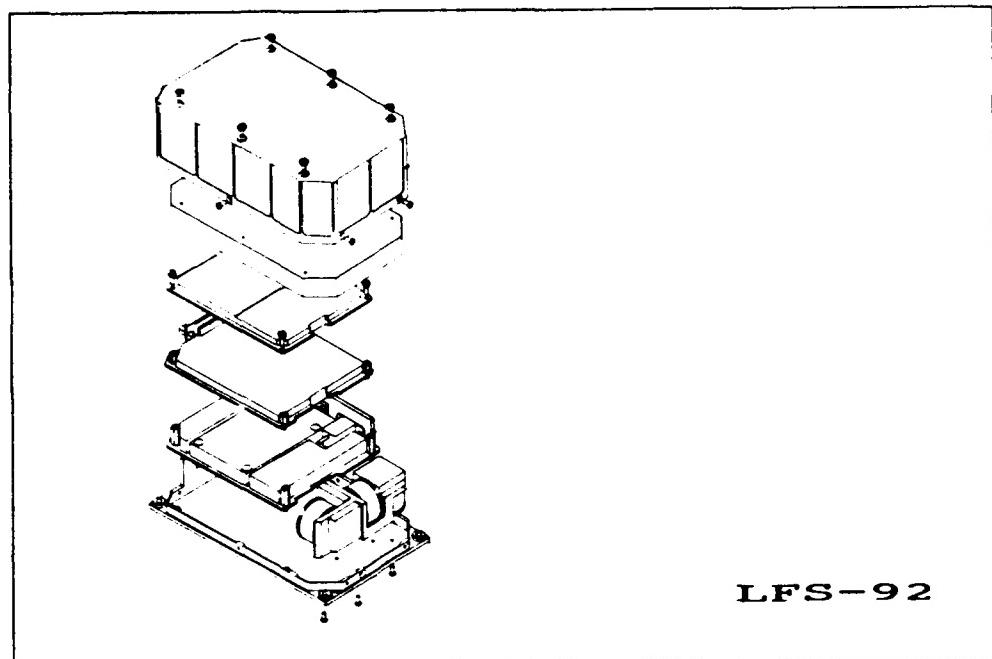
- trainer aircraft
- utility aircraft
- business turboprop
- commuter aircraft
- lightweight helicopters

The digital output and optional synchro output meet the requirements of modern digital avionic systems as well as conventional analogue installations.

References

- 1) Büschelberger, Schröder "Packaged Fiber Optic Gyros"
AGARD Conference Proceedings No.435
Guidance and Control of Precision Guided Weapons
Geilo, Norway May 1988
- 2) Handrich, Krings, Büschelberger
"Flight Demonstration of Fiber Optic Gyros AHRS"
Conference Proceedings Symposium Gyro Technology
Stuttgart, September 1989
- 3) Büschelberger, Kemmler "Closed Loop Fiber Optic Gyro
Triad" SPIE Conference Proceedings Vol.1585
15th Anniversary Conference on Fiber Optic Gyros
Boston, September 1991
- 4) Handrich, Büschelberger, Kemmler, Krings
"LFS-90 - A Modular System Design with Fiber Optic
Gyros" Conf. Proc. Symposium Gyro Technology
Stuttgart, September 1991
- 5) Buitkamp "Drehmessungen ohne Achse und Teilkreis"
ROBOTER, September 1991, p 26-28





ACCELEROMETRE SILICIUM POUR A.H.R.S. ET SYSTEMES HYBRIDES DE NAVIGATION

J. LECLERC A. DEFOSSE O. LEFORT

SEXTANT AVIONIQUE
25, Rue Jules Védrines
26027 VALENCE Cédex
FRANCE

RESUME

La société SEXTANT Avionique possède une large gamme d'accéléromètres pour les différentes applications de Navigation. Depuis 1980, un souci constant d'adéquation au besoin a conduit SEXTANT Avionique à développer des capteurs micro-usinés et en particulier des accéléromètres. Plusieurs types de matériaux ont été utilisés et nous exposons les raisons du choix du silicium pour les accéléromètres destinés aux AHRS. L'architecture de l'accéléromètre, les technologies utilisées, les modes de fonctionnement sont décrits. Ensuite sont développées les performances actuelles des différents types d'accéléromètres silicium, ainsi que leur domaine d'application. En conclusion, les perspectives à plus long terme sont évoquées.

LISTE DES SYMBOLES

g	: Accélération de la pesanteur (9,8 m/s ²)
E	: Module d'élasticité d'un matériau (Young) (N/mm ²)
$\gamma_{x,y,z}$: Accélérations linéaires suivant les trois axes de l'accéléromètre(m/s ²)
$\Omega_{x,y,z}$: Vitesse angulaire suivant ces mêmes axes (rad/s)
ϵ_0	: Permittivité diélectrique du vide (F / m)
d_0	: Distance moyenne inter-electrodes (μm)

INTRODUCTION

La société SEXTANT Avionique possède une large gamme d'accéléromètres pour les différentes applications de Navigation et de Conduite du vol.

Elle fabrique un nouveau type de centrales de navigation et de pilotage, les AHRS " Attitude and Heading Reference Systems " qui utilise de manière optimale les différentes mesures de position et de mouvement d'un mobile dans l'espace, accéléromètres, gyromètres, anémomètres, magnétomètres ainsi que les mesures de position optiques ou radioélectriques ; Radar DOPPLER, Récepteur OMEGA, G.P.S.

Ces centrales profitent du savoir faire spécifique de la société dans les domaines de l'anémobarométrie, la magnétométrie et les techniques de centrales inertielles liées

L'application de ce concept aux équipements de l'an 2000 se doit d'intégrer les technologies de réalisation de capteurs les plus prometteuses .

C'est ainsi que pour le système de référence de navigation de l'hélicoptère TIGRE, dans un souci constant d'adéquation au besoin, ont été développés le Gyrolaser Triaxial PIXYZ et le Micro ACCéléromètre Silicium MACSI®.

Ce dernier, qui sera l'objet de cette présentation, est particulièrement représentatif de l'évolution des micro-technologies appliquées à la réalisation de capteurs . Il a été développé avec l'aide du C.S.E.M. (Centre Suisse d'Electronique et de Microtechniques).

DESCRIPTION DU BESOIN

L'AHRS fournit principalement les informations de cap et d'attitude nécessaires au pilotage et à la navigation et les informations de vitesse angulaires et d'accélération pour le pilote automatique .

-Un ensemble capteurs, composé de capteurs gyrométriques et accélérométriques permet de mesurer les vitesses angulaires et les accélérations du porteur.

-Un ensemble algorithmique AHRS, basé sur celui d'une Plateforme Inertie Pure, réalise les traitements concernant la plate-forme virtuelle (calcul des angles d'attitude par intégration des vitesses angulaires, calcul des vitesses et positions par intégration des accélérations). Cependant, les techniques d'hybridation de la plateforme avec des informations complémentaires permettent d'utiliser des capteurs plus simples et de limiter la dégradation des performances dans le temps.

- Hybridation en cap magnétique avec le magnétomètre
- Hybridation en vitesse air ou sol par anémométrie ou radar Doppler
- Hybridation en position Omega ou GPS.

Cette ouverture au niveau des caractéristiques métrologiques a amené SEXTANT Avionique à reconsidérer la fonction accélérométrie . Les accéléromètres, moins performants que les accéléromètres de classe inertie, devaient toutefois être moins encombrants et moins coûteux .

Les caractéristiques de ceux-ci sont brièvement décrites ci-après.

Etiende de Mesure	: 4 à 20 g
Erreur de biais :	: 10^{-3} E.M.
Erreur de facteur d'échelle :	: $5 \cdot 10^{-4}$

Elles devront être garanties dans tout le domaine de vibration, température et dans le temps.

L'optimisation des coûts a conduit SEXTANT Avionique à considérer les micro-technologies pour développer un nouveau type d'accéléromètre.

LES MICRO-TECHNOLOGIES

Les micro-technologies ont rendu possible la conception et réalisation de détecteurs très petits qui utilisent les excellentes propriétés mécaniques du silicium ou du quartz. -Encombrement réduit grâce aux techniques de la microélectronique ; Les procédés de gravure anisotropique suivant les directions cristallines bien définies associées aux techniques de photolithographie permettent la réalisation collective de motifs mécaniques complexes et de faible dimension.

- Grâce à leur constitution mono-cristalline, le quartz et le silicium présentent une hystérésis faible et une stabilité mécanique élevée.

- Intégration du traitement du signal sur un même hybride ou une même puce car le détecteur et son électronique sont de même nature ou nécessitent les mêmes techniques de connexion.

- Disponibilité des matières premières et maturité des procédés technologiques dérivés des technologies de la microélectronique.

- Et surtout, réduction drastique des coûts grâce aux techniques d'usinage collectif.

CHOIX DU MATERIAU

Le quartz est un matériau traditionnellement utilisé pour ses qualités mécaniques et de stabilité (résonateurs, références de fréquence, tube de Bourdon). Il est donc particulièrement adapté à la réalisation de capteurs.

Le silicium a toutefois des propriétés intrinsèques; module d'élasticité, dérive en température et stabilité équivalents à ceux du quartz.

Leurs principales caractéristiques sont regroupées dans le tableau ci-dessous. Les stabilities mécaniques du quartz et du silicium y sont comparées par l'intermédiaire de la stabilité en fréquence dans le temps de résonateurs réalisés dans ces mêmes matériaux.

	E (N/mm ²)	$dE/E \times dT$ (K ⁻¹)	Df/f
SILICIUM	1,66E 11	-85E-6	3E-5
QUARTZ	0,9E 11	-47E-6	≈ 0

Les qualités mécaniques du quartz peuvent cependant, suivant son utilisation, être affectées par les dépôts métalliques rendus nécessaires par le fait qu'il est isolant. Ses propriétés piézo-électriques peuvent être un atout pour la conception de capteurs à résonateur.

Le silicium permet d'utiliser des matériaux et outils de fabrication communs à ceux de la microélectronique donc moins chers, et des technologies de masquage plus simples que pour le Quartz.

L'impossibilité de définir des arrêts de gravure dans le quartz rend le contrôle de l'épaisseur de certains profils (membranes, bras de suspension) imprécis car seul contrôlé par le temps de gravure.

De plus, la possibilité d'intégration de l'électronique de traitement du signal sur le détecteur silicium ne peut être écartée.

En conclusion, le quartz et le silicium sont deux matériaux utilisables, par leurs caractéristiques mécaniques, par ailleurs assez voisines, pour la réalisation de capteurs.

En fait, le choix du matériau influe beaucoup sur la conception du détecteur. Ainsi, ce sont le domaine d'application et les différentes contraintes techniques et économiques qui imposent un argument plutôt qu'un autre et, finalement, le choix de la filière technologique pour la réalisation du capteur.

CONCEPTION DU DETECTEUR

Le silicium permet globalement, dans le cas qui nous préoccupe, d'envisager des coûts de production moins élevés.

La structure est construite par assemblage de cinq plaques de silicium et de verre par soudure anodique sous pression et nature de gaz contrôlées. Le verre assure l'isolation électrique, le silicium, légèrement dopé, les fonctions mécaniques et électriques.

La tranche de silicium centrale est usinée par gravure anisotropique dans une solution aqueuse de potasse pour définir une masse sismique maintenue par des bras de suspension.

A proximité des deux faces de la masse sismique, des électrodes sont déposées sur le verre et définissent ainsi deux condensateurs dont les variations suivant la position de la masse sismique sont opposées, permettant la détection de la position de la masse.

La figure 1 montre un détecteur éclaté et sa masse sismique.

La masse sismique peut être suspendue de deux manières différentes définissant deux versions de détecteurs dont la sensibilité pourra être ensuite, adaptée au besoin en changeant la section des bras.

Les deux configurations sont illustrées en Fig. 2 et 3. L'analyse a été menée avec différentes étendues de mesures pour les versions "Deux bras" et "Quatre bras".

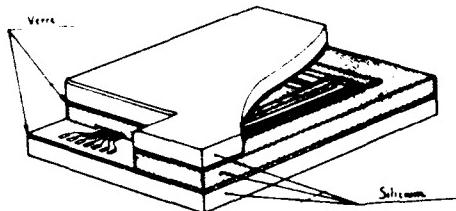


Fig. 1 Détecteur MACSI (R) : éclaté

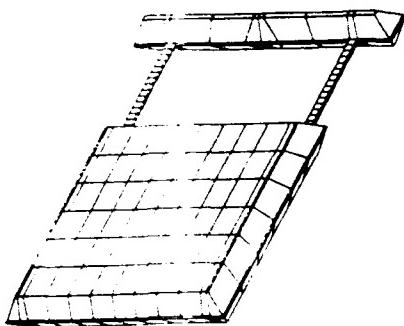


Fig. 2 Version "Deux bras"

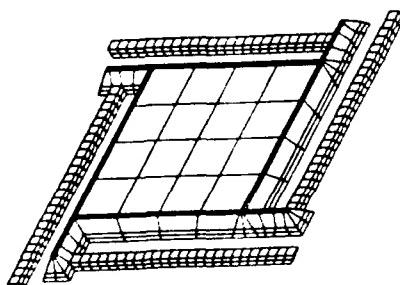


Fig. 3 Version "Quatre bras"

Le signal utile est la différence de capacités $\delta C = C_1 - C_2$ entre la masse sismique et les deux électrodes (Fig. 4). Il s'exprime comme une fonction de l'accélération appliquée au détecteur suivant les axes X, Y, Z, et les rotations autour de ces mêmes axes, en l'absence de forces électrostatiques.

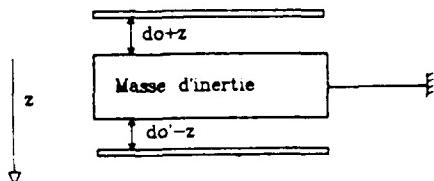


Fig. 4

Pour le modèle "Deux bras"

$$\delta C = K [\gamma_x + \gamma_0 + k_1 \gamma_x + k_2 \gamma_x^2 + k_3 \gamma_x^3 + k_z \gamma_z + a_1 \Omega_x + a_2 \Omega_y \Omega_z]$$

Pour le modèle "Quatre bras"

$$\delta C = K [\gamma_x + \gamma_0 + k_2 \gamma_x^2 + k_3 \gamma_x^3 + a_1 (\Omega_x + \Omega_z) + a_1 (\Omega_x^2 + \Omega_z^2) + a_2 [(\Omega_x \Omega_y)^2 + (\Omega_x \Omega_z)^2]]$$

avec $\gamma_x, \gamma_y, \gamma_z$, les accélérations suivant l'axe sensible, et les axes perpendiculaires
 $\Omega_x, \Omega_y, \Omega_z$, les vitesses de rotation suivant ces axes .

Dans le tableau Tab. 1 sont comparées les valeurs de ces différents coefficients .

(Caractéristiques en boucle ouverte)

Etendue de mesure	0,1 g	0,1 g	30 g	30 g
Nombre de bras	2	4	2	4
K (pF/g)	26	26	0,085	0,085
k _y (-)	6 E-4	0,08	6 E-4	0
k ₂ (g-1)	0,18	0,08	5,8E-4	2,6E-4
k ₃ (g-3)	1,55	1,15	1,6E-5	1,25E-5
k _{zy} (g-1)	1,3E-3	0	4,4E-6	0
a _x (μg/rad/s ²)	22	0,15	22	5E-4
a _y (μg/rad/s ²)	0	0,15	0	5E-4
a ₁ (μg/(rad/s) ²)	0	2,2E-5	0	7E-8
a _{yz} (μg/(rad/s) ²)	21	0	21	0
a ₂ (μg/(rad/s) ⁴)	0	2,2E-5	0	7E-8

Tab. 1 comportement des deux détecteurs

La version quatre bras se caractérise par un couplage d'axe k_{zy} inexistant et une sensibilité à la rotation moindre . Ceci définit les caractéristiques mécaniques du détecteur envisagé. Les performances calculées sur ces deux types de détecteur ne permettent toutefois pas d'envisager leur utilisation pour des applications accéléro AHRS.

L'asservissement de la masse à déviation nulle par une boucle de position amoindrit l'influence des défauts et dérives mécaniques et apporte une amélioration importante de la précision du capteur.

La polarisation de ces électrodes permet, par ailleurs, la création de forces électrostatiques assurant, avec l'aide d'une électronique d'asservissement, le maintien de la masse à position fixe, sans dispositif supplémentaire comme un circuit magnétique et des aimants.

$$F = K \left[(V_1^2 - V_2^2) + (V_1^2 - V_2^2) \frac{2z}{d_c} \right], \quad K = \frac{\epsilon_0 S}{2 d_c^2}$$

avec V_1 et V_2 les tensions entre les électrodes et la masse sismique.

Le terme $K(V_1^2 - V_2^2)$ représente la force de rappel exercée par l'asservissement et le terme $K(V_1^2 - V_2^2) \times 2z/d_c$ peut être assimilé à une raideur de telle sorte que la raideur mécanique du système puisse être réglée. En effet, la fréquence de résonance du système $f = 1/2\pi\sqrt{K_{tot}/M}$ peut être ajustée pour modifier la fonction de transfert dynamique de l'accéléromètre asservi.

DESCRIPTION DU CAPTEUR

Version non asservie :

En fonctionnement en boucle ouverte, la sensibilité de l'accéléromètre est définie par la raideur des bras de suspension et dépend du déplacement attendu de la masse soit, indirectement, de la linéarité escomptée. Les déplacements de la masse sismique sont mesurés au point milieu du pont capacatif défini par le détecteur. (Voir Fig. 5).

Version asservie :

En fonctionnement asservi, des tensions V_0+Vs et V_0-Vs sont appliquées aux deux électrodes. Ce décalage V_0 permet de linéariser la réponse du capteur :

$$Vs = 4KV_0/M \times \gamma$$

L'information de sortie du capteur est ainsi directement accessible sous forme linéaire analogique amplifiée facilement utilisable par le calculateur sans circuit spécifique complexe d'analyse de signaux.

L'asservissement permet par ailleurs la réduction des erreurs dues aux imperfections mécaniques. Le déplacement de la masse sismique devient négligeable et la non linéarité du capteur, donc les erreurs de rectification, faibles. Le couplage d'axe k_{zy} est, par ailleurs, annulé. La bande passante du capteur est facilement ajustée en intégrant un filtre dans la boucle d'asservissement. (voir Fig. 6.)

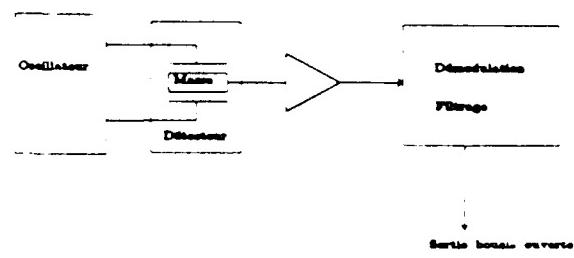


Fig. 5 Capteur non asservi

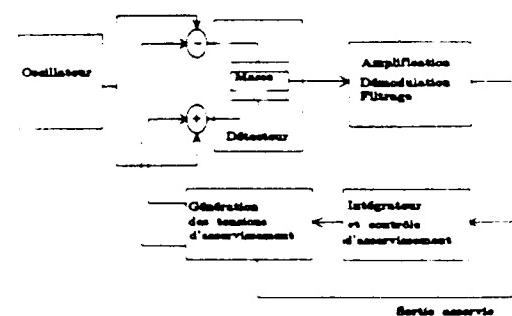


Fig 6. Capteur asservi

REALISATIONS ET RESULTATS OBTENUS

Une version asservie de MACSI® est en cours de développement, pour le besoin A.H.R.S pour application hélicoptère, porteur dont l'environnement vibratoire est particulièrement sévère.

Cet accéléromètre réalise un compromis entre un comportement très amorti, permettant d'éviter les erreurs de rectification et une bande passante élevée. Ce problème est particulièrement important car le niveau de vibration du porteur est bien supérieur à celui de l'accélération statique. Par ailleurs, les problèmes de packaging, report de détecteur sur substrat céramique, précision de positionnement et stabilité de la liaison, stabilité devant la pression ambiante et surtout tenue et stabilité en température ont été optimisés pour cette application. Les résultats obtenus sont résumés par les courbes de précision (fig. 7) et stabilité court terme (fig. 8).

CONCLUSION

Une famille de micro-acéléromètres silicium MACSI® a été décrite. Il apparaît que les micro-technologies permettent d'atteindre des performances compatibles des applications "Attitude and Heading Reference Systems" AHRS, et des coûts de réalisation très attractifs. Le domaine d'application de ces micro-acéléromètres devrait s'élargir et les conditions de succès de ces technologies seront renforcées à chaque fois que l'environnement d'utilisation sera sévère, les quantités nécessaires et les contraintes économiques seront importantes.

REFERENCES

- F RUDOLF A micromechanical capacitive accelerometer with a two point inertial-mass suspension, Sensors and Actuators, 4 (1983) 191-198
- K. E. PETERSEN Silicon as a mechanical material, Proceedings of the IEEE, vol 7, No. 5, May 1982
- O. LEPORT A Miniature, Low Cost, Silicon Micromachined Servo-Accelerometer, Proceedings of the Symposium Gyro-Technologies (1988) Stuttgart, Germany
- M. DUFOUR A comparison between micromachined pressure sensors using quartz or silicon vibrating beams, TRANSDUCERS 91 Digest of technical papers (p 668-671), San Francisco.

ERREUR DE LINEARITE ET HYSTERESIS CALIBRAGE EN TEMPERATURE

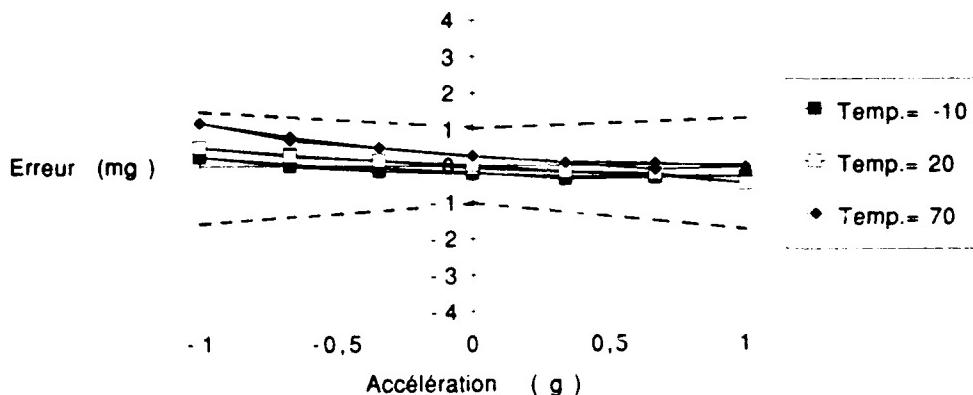


Fig. 7 Précision dans le domaine de température

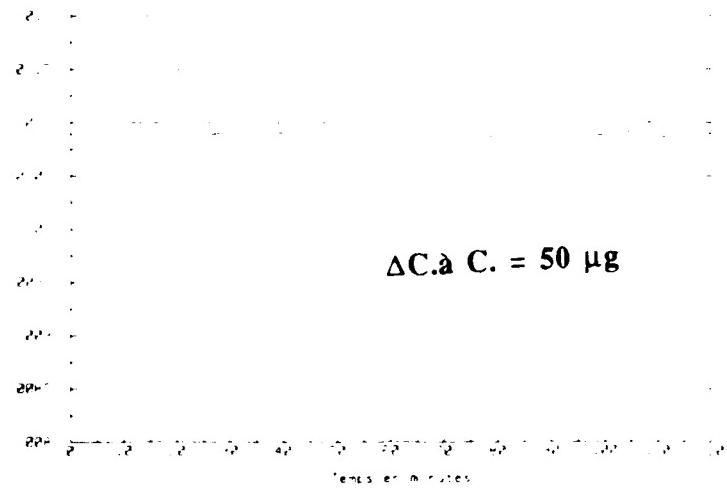


Fig. 8 Stabilité courante du capteur

INTEGRATION OPTIMISEE DE L'INERTIE ET DU GPS

par

**M. Loïc Camberlein, Mme Bernadette Capit
et M. Pascal Debanne**

SAGEM
6, Avenue d'Iena
75016 Paris, France

RESUME

ULISS NG est une centrale de navigation Inertie-GPS, concrétisant la synergie entre Inertie et GPS et aboutissant à un bon compromis coût performances. Elle appartient à la famille ULISS qui équipe déjà différents types d'avions. Son récepteur GPS intégré, miniaturisé douze canaux code P est composé d'un module radio fréquence et d'un module de traitement numérique et représente au total 0.7 l, 0.7 kg et 12 W. La poursuite en parallèle des satellites conduit à une précision, une continuité et un lissage maximum de la navigation, tout en minimisant le temps de réaction et la taille de l'horloge.

Le couplage serré Inertie/GPS, c'est-à-dire une aide inertiale directe, associée à un filtre serré de navigation (Kalman), a été conçu pour optimiser les capacités opérationnelles, les performances, ainsi que la stabilité de la solution de navigation.

Quatre versions du récepteur GPS embarqué ont été développées : C/A-SPS, C/A-PPS et P(Y)L1 et P(Y)L1'L2.

C'est la version code P(Y) bi-fréquence qui est présentée dans cet article.

Des essais en vol, sur chasseur Mirage III, vont être effectués au Centre d'Essais en Vol de Brétigny (CEV) à partir du mois de mai.

ABSTRACT

ULISS NG is an Inertia-GPS unit and an interesting example of INS/GPS synergy and performance/cost trade-off. It belongs to the ULISS family of INS already fitted to a number of different aircrafts. Its miniaturized code twelve channel embedded receiver, with its RF and digital modules, totals 0.7 l, 0.7 kg and 12 W. Its parallel tracking maximizes navigation continuity, smoothness, accuracy and minimizes reaction time and oscillator size.

The tight Inertia-GPS coupling, i.e. tight inertial aiding and tight hybrid navigation Kalman filter, has been designed to favour operational capabilities, performance and solution stability.

Four versions of the SAGEM embedded GPS receiver have been developed C/A-SPS, C/A-SA-PPS and P(Y)L1 and P(Y)L1'L2. The P code version is presented in the paper.

Flight tests are being conducted on a Mirage III at the French Official Flight Test Center of Brétigny starting in May.

1 - INTRODUCTION

Bien des choses ont été écrites sur l'exceptionnelle complémentarité du GPS et de l'Inertie, sur les multiples niveaux possibles d'intégration matériel logiciel et sur le compromis correspondant entre l'efficacité de la synergie et le rapport performances/coût. Ainsi, nous n'entrerons pas dans le détail de ces sujets, par ailleurs

développés dans de nombreuses publications (cf. références [2], [3], [6], [8], [12]).

En résumé, cette complémentarité exceptionnelle vient, d'une part, de la très bonne précision à long terme du GPS (position tridimensionnelle et vitesse) et, d'autre part, de la grande bande passante et de la très bonne précision à court terme de l'Inertie (position tridimensionnelle, vitesse et accélération). Elle provient aussi de leur égale et remarquable couverture de toute la planète, 24 heures sur 24, grâce à l'autosuffisance et l'immunité au brouillage de l'Inertie (contrairement au GPS) et grâce aux informations de temps données par GPS et d'attitude données par l'Inertie.

Les différents niveaux possibles d'intégration matérielle ou logiciels vont de la fourniture d'équipements séparés, GPS et centrale inertuelle, à une complète intégration dans un ensemble unique où un seul filtre de Kalman adaptatif peut contrôler à la fois les canaux GPS et la navigation hybride.

Entre ces deux extrêmes, la synergie va d'un minimum à un maximum, bien que cette dernière limite, pour idéale qu'elle soit, puisse conduire à une solution pratiquement irréalisable.

Le rapport performances/coût et le compromis sur l'efficacité de la synergie se porte généralement sur :

- les performances, c'est-à-dire la précision, la résistance au brouillage, les capacités d'évolutions dynamiques (problème de masquage d'antenne), la dynamique maximale (facteur de charge), les interruptions GPS, les possibilités d'alignement inertiel rapide en vol ou à la mer,
- les contraintes d'installation dans un aéronef : le nombre d'équipements, les encombrements, le poids, la consommation, la charge bus ainsi que la plus ou moins grande facilité de validation et de test,
- le contrôle d'intégrité, la fiabilité, la robustesse, la redondance, la stabilité de la solution de navigation hybride.

Cet exposé a pour objet de présenter la centrale Inertie/GPS ULISS NG, qui est un exemple intéressant de l'efficacité de la synergie et d'un excellent rapport performances/coût. Les sous-ensembles Inertie et GPS sont décrits en détail dans les paragraphes 2 et 3. Le paragraphe 4 présente le couplage serré choisi pour tirer le meilleur parti possible de la synergie INS/GPS. Le paragraphe 5 traite des essais en vol sur avion de combat Mirage III.

2 - "TOUT EN UN" : INTEGRATION PHYSIQUE INERTIE-GPS

2.1 NOUVELLE OPTION DANS LA FAMILLE ULISS

La centrale Inertie/GPS ULISS NG (figure 1) présente l'option GPS au sein de la famille des centrales inertielles ULISS. La technologie ULISS a été conçue afin de répondre à une gamme d'applications allant d'une simple centrale inertuelle de navigation à celle

d'une centrale de navigation-attaque. Sa conception a abouti à d'excellentes performances inertielles, une fiabilité et une maintenabilité remarquable, ainsi qu'une grande facilité d'intégration fonctionnelle et une souplesse d'adaptation et d'installation.

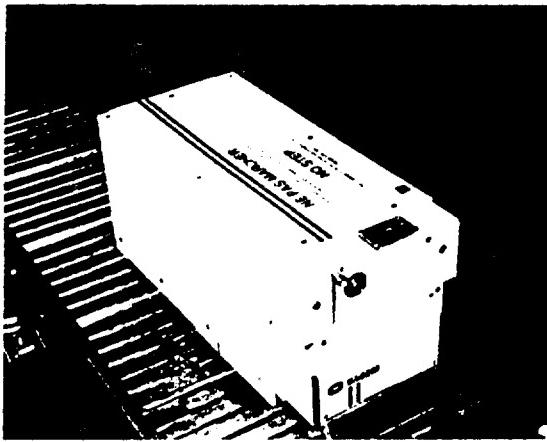


figure 1. Système inertiel/GPS ULISS NG

La même technologie de haut niveau d'intégration est utilisée par les différentes versions de la famille ULISS, dont 1200 exemplaires sont actuellement en service sur quatorze types d'avions différents dans douze forces aériennes et aéronavales. Quelle que soit la version, le format (3/4 ATR) et le poids (16 kg) restent identiques. Le cœur inertiel et plus de 80 % des modules électroniques sont communs à toute la famille. La technologie ULISS s'appuie sur une conception modulaire nettement orientée logicielle, sur l'état de l'art en matière d'inertie, sur des options d'entrées/sorties couvrant l'ensemble des applications et sur des réserves pour des options futures. Le bloc diagramme fonctionnel d'ULISS NG est donné figure 2. Dans le même équipement léger et le faible volume, sont rassemblés :

- le récepteur intégré GPS (la nouvelle option ULISS),
- le cœur inertiel constitué d'une plate-forme à cardans, miniature et performante, utilisant l'état de l'art en matière de gyroscopes secs accordés (S 040) et d'accéléromètres secs (A 310),
- l'électronique du cœur utilisant des composants VLSI et micro-électroniques,
- le calculateur de navigation, d'une puissance de 1 MIPS, dont les fonctions mémoire et calcul sont disponibles pour des fonctions autres que la navigation inertuelle,
- l'électronique d'entrée/sortie capable de différents interfaces standards telles que des bus multiplexés à 1 Mbit/s, type 1553 ou Digibus, bus ARINC et des entrées/sorties analogiques ou discrètes,
- l'alimentation des fonctions précédentes.

La figure 2 montre aussi les différentes sorties disponibles : inertie pure, GPS pur et aidé, navigation hybride.

ULISS NG est équipé d'un câble coaxial de liaison avec l'antenne GPS qui peut être soit une antenne fixe avec son préamplificateur faible bruit ou une antenne adaptative avec son module électronique de contrôle. Dans l'électronique du système, les deux modules RF et traitement numérique du récepteur GPS sont semblables aux autres modules électroniques de la

centrale et ne représentent que 15 % du volume électronique total. Ils sont décrits au paragraphe 3.

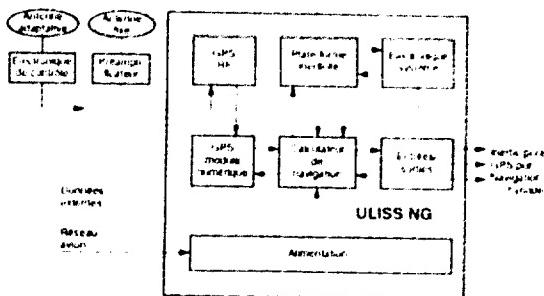


figure 2. Bloc diagramme fonctionnel simplifié de l'ULISS NG

2.2 UN EQUIPEMENT "TOUT EN UN"

ULISS NG démontre, par sa taille et son faible poids, que le concept "Tout en un" permet d'obtenir une synergie matérielle INS/GPS remarquable, sans comparaison avec une solution à ensembles séparés. La réduction dans un rapport 2 du nombre des équipements apporte presque le même taux de réduction drastique pour les volumes, le poids, la consommation et la charge bus, et ceci présente un intérêt tout particulier pour les aéronefs de haute performance.

Dans le cas des modernisations d'avions, l'installation de l'option GPS devient quasiment "indolore" (exception faite de l'installation antenne qui est à faire dans tous les cas) dans la mesure où l'encombrement, avec ou sans GPS, est le même.

La validation globale du système avion le test et les échanges d'informations sur le bus sont grandement simplifiés

3 - UN RECEPTEUR EMBARQUE "TOUT EN VUE"

3.1 CRITERES GENERAUX DE CONCEPTION

L'architecture du récepteur, illustrée figure 3, est issue de la prise en compte des besoins utilisateurs (performance, intégrité et flexibilité dans un environnement de dynamique et de brouillage sévère) et de l'utilisation d'une technologie moderne très intégrée. Les points forts sont les suivants :

• Poursuite parallèle "Tout en vue".

L'expression maintenant familière de "tout en vue" désigne la possibilité du récepteur GPS de poursuivre tous les satellites visibles. Avec le déploiement de la constellation de vingt et un satellites et plus, jusqu'à onze satellites pourront être visibles en un point donné. En conséquence, il est souhaitable de disposer d'autant de canaux physiques indépendants que de satellites visibles afin de bénéficier au maximum des informations disponibles.

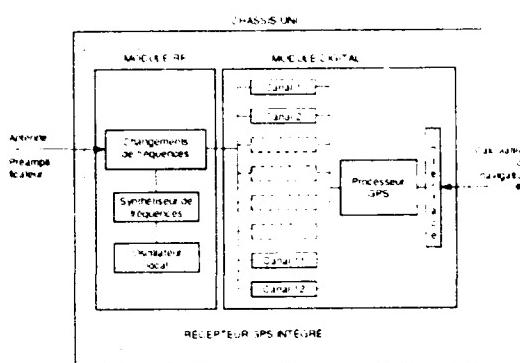


figure 3 Bloc diagramme simplifié du récepteur GPS intégré Code P (douze canaux parallèles)

Le concept "Tout en vue" maximalise :

- la continuité et le lissage des sorties de navigation (en effet, dans les applications aéroportées, les satellites sont fréquemment masqués en raison de l'inclinaison de l'avion dans les virages et les manœuvres).
- la précision des sorties de navigation qui augmente avec le nombre de satellites accrochés.
- la fiabilité et l'intégrité des sorties de navigation, dans la mesure où il réduit le temps de détection de satellites en panne et autorise une redondance matérielle.

La poursuite de tous les satellites visibles se fait par des canaux physiques, indépendants et parallèles. Il n'est plus aujourd'hui besoin de démontrer les possibilités bien supérieures de la poursuite et du traitement parallèle synchrones un traitement séquentiel ou par rapport à même un multiplexage rapide, en termes de réjection des interférences, de dynamique et de robustesse. La poursuite parallèle minimise aussi le temps de réaction à la mise sous-tension - sans qu'il soit nécessaire de disposer d'informations externes (position ou temps) - grâce à la mise en œuvre d'une recherche initiale optimisée et l'absence de délai de préchauffage de l'oscillateur local. Le besoin de stabilité à court terme de l'oscillateur local est moindre, grâce au traitement parallèle qui permet en outre de choisir un oscillateur de dimension réduite.

• Miniaturisation et traitement presque entièrement digital.

L'état de l'art de la technologie permet un traitement du signal GPS presque entièrement digital. C'est ainsi qu'après la conversion analogique-numérique dans le module RF, tout se fait en numérique, grâce à l'utilisation de circuits ASIC et de "VLSI".

Cette solution a des avantages évidents

- une grande miniaturisation du récepteur GPS, ce qui représente un plus pour l'intégration INS/GPS.
- l'optimisation du rapport signal à bruit,
- un fonctionnement indépendant des conditions de température et du temps (contrairement à l'analogique).
- une grande fiabilité.

- une grande souplesse de développement et d'évolution, tout étant réglé et contrôlé par logiciel.

La poursuite parallèle "Tout en vue" et la très grande miniaturisation résultant de l'utilisation d'une technologie numérique très intégrée ont permis d'obtenir d'excellentes performances avec une conception relativement simple dans un volume réduit. Ceci n'aurait sans doute pas été réalisable autrement.

3.2 DESCRIPTION MATERIELLE

Le récepteur GPS (figure 4) se compose de deux modules différents : un module radio fréquence bifréquence et un module digital.

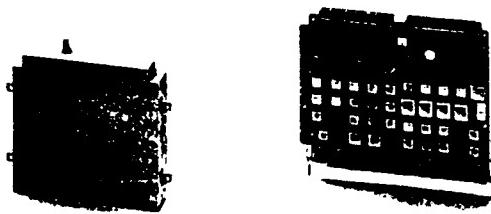


figure 4 - Récepteur GPS "Tout en vue" intégré
(0,71-0,7 kg - 12 W)

Le module radio fréquence (125 × 110 × 30 mm) contient l'étage de réception (RF) et le synthétiseur de fréquences. L'étage de réception effectue pour chacune des voies L1 C/A, L1 (P), L2 (P) les changements de fréquence nécessaires, l'amplification et le filtrage jusqu'à la fréquence de base. Les signaux sont convertis en numérique à la sortie du module RF. Le synthétiseur délivre, à partir d'un oscillateur local, toutes les fréquences et les signaux d'horloge nécessaires au traitement des signaux GPS. Un module RF bifréquence, d'une dimension si petite, n'a pu être réalisé qu'en utilisant les technologies RF de pointe, telles que des amplificateurs FET faible bruit à AsGa, les filtres acoustiques à ondes de surface et la technologie ECL en "VLSI". Ces composants électroniques ne contribuent pas seulement à réduire la consommation, mais permettent aussi d'optimiser le rapport signal à bruit.

Le module digital (118 × 156 × 10 mm) est double face et comporte 24 canaux parallèles. Un processeur unique est dédié au traitement du signal, au contrôle des boucles et au traitement de la navigation GPS. Les transferts de données internes entre GPS et Inertie se font au travers d'une interface DMA directe qui minimise les retards et autorise une datation précise qui aurait été difficile à réaliser avec des équipements GPS et Inertie séparés. Le choix d'ASIC VLSI en technologie CMOS et d'un processeur de signal numérique suffisamment puissant ont aussi contribué à la miniaturisation de l'ensemble et à une faible consommation. Les caractéristiques physiques obtenues pour ce récepteur GPS intégré code P sont :

- volume 0,7 l.
- poids 0,7 kg.
- consommation 12 W.

Ces caractéristiques sont celles de la version la plus performante P(Y) L1/L2 des récepteurs GPS intégrés développés par SAGEM.

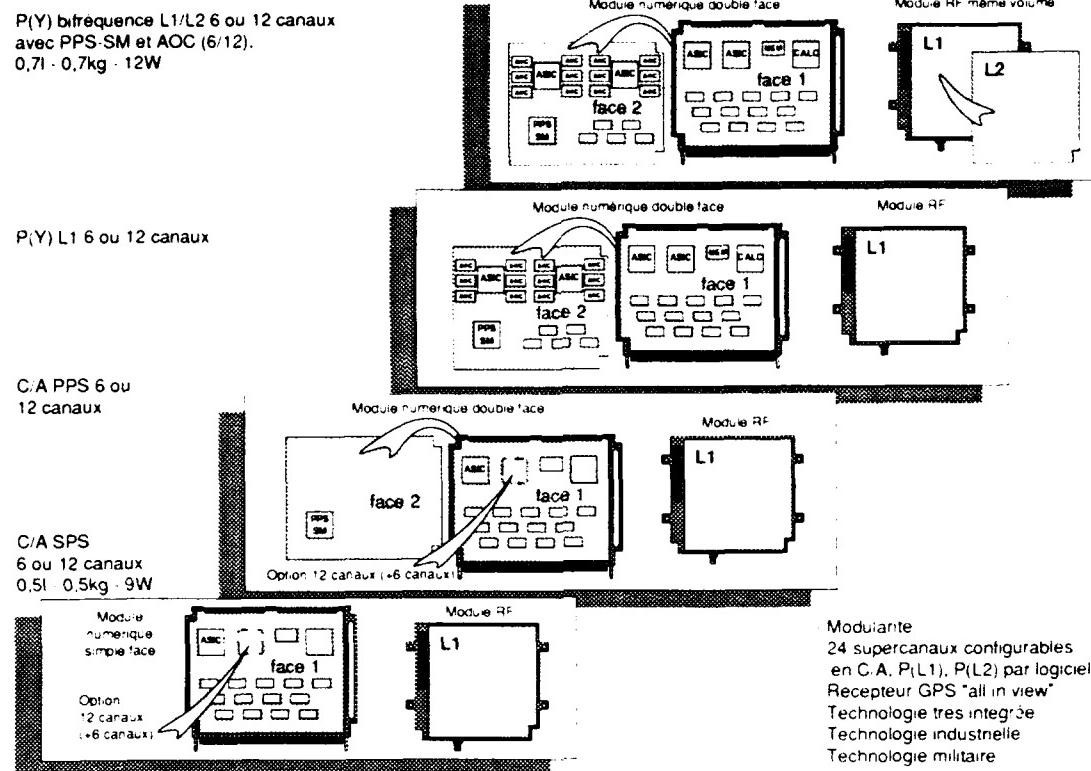


figure 5 - Gamme des récepteurs GPS militaires SAGEM

Le choix de la modularité ayant été fait dès la conception du récepteur, une gamme complète de récepteurs intégrés est dérivée de la version la plus sophistiquée par simple dépeuplement des modules ainsi que le montre la figure 5 :

- récepteur C/A 6 ou 12 canaux SPS.
- récepteur C/A 6 ou 12 canaux PPS.
- récepteur P (Y) L1 6 ou 12 canaux.
- récepteur P (Y) L1/L2 6 ou 12 canaux.

Une des originalités réside dans le fait que les canaux des ASIC sont configurables entièrement par logiciel en C/A L1, P L1 ou P L2.

3.3 DESCRIPTION LOGICIELLE

Le logiciel est de conception modulaire. Toutes les fonctions liées au GPS sont réalisées par un processeur unique. Le logiciel GPS, organisé autour d'un moniteur temps réel, comprend les fonctions suivantes : logiciel de contrôle canal, logiciel de gestion des canaux et des satellites, logiciel de navigation et enfin logiciel d'interface. Toutes ces fonctions sont elles-mêmes divisées en sous-fonctions qui peuvent être implémentées en différentes tâches temps réel, suivant le besoin.

- 1 - Le logiciel canal est le cœur du traitement de signal GPS. Il commande les boucles de code et de porteuse afin de les verrouiller sur les satellites.

suivant le mode choisi par le moniteur d'accrochage ; ceci terminé, il gère ou mesure la phase de code, le Doppler et le Doppler intégré, la phase de porteuse et les marqueurs temporels associés, les indicateurs de boucles et, enfin, il collecte les informations de navigation satellites.

- 2 - Le logiciel de gestion des canaux et des satellites est divisé en trois sous-fonctions : gestion satellites, gestion canal et moniteur d'accrochage. Il procède à l'affectation, de la meilleure manière possible, des satellites aux canaux disponibles, en fonction de la procédure en cours dans le récepteur (initialisation, navigation, réacquisition après masquage). Il élabore et envoie, à chaque canal, les paramètres nécessaires à l'accrochage d'un satellite donné, en fonction de son état antérieur et des informations disponibles pour calculer les paramètres d'accrochage. Ce logiciel bénéficie pleinement du grand nombre de canaux.

Une grande souplesse dans la gestion des canaux est apportée par le fait que les canaux sont configurables entièrement en logiciel en C/A, P L1 ou P L2.

- 3 - Le logiciel de navigation possède trois sous-fonctions principales : le traitement des mesures GPS, le décodage des informations de navigation et leur traitement et le calcul de la position GPS, de la vitesse et des paramètres spécifiques (GDOP, EPE, etc ...).

La première sous-fonction élaboré une "pseudo-distance" pour chaque satellite, à partir de la distance mesurée en fractions de millisecondes, du nombre entier de millisecondes et des différentes corrections à appliquer à chaque mesure : erreur ionosphérique mesurée, correction troposphérique,

relativiste, offsets d'horloge du satellite ou du récepteur.

La seconde sous-fonction est un décodage classique des informations du message de navigation, tel qu'il est décrit dans le document ICD 200, en particulier des éphémérides, de l'almanach, des paramètres d'horloge satellites, des différents indicateurs de validité, des drapeaux et des indicateurs de qualité qui peuvent influencer la navigation ou le mode d'accrochage du récepteur.

La troisième sous-fonction calcule la solution purement GPS en position, vitesse et offset d'horloge, à partir de toutes les informations disponibles sur l'ensemble des canaux. L'amélioration de la précision de position, à mesure que le nombre de satellites acquis augmente, est très significative. Cette sous-fonction élaborer aussi le facteur géométrique de la constellation utilisée ainsi qu'un indicateur de qualité de la solution GPS.

- 4 - Le logiciel d'interface gère les échanges de données avec le calculateur de navigation de la centrale inertie.

4 - COUPLAGE SERRE INERTIE-GPS

4.1 DESCRIPTION GENERALE

Le schéma fonctionnel général du logiciel de l'ULISS NG est donné en figure 6 :

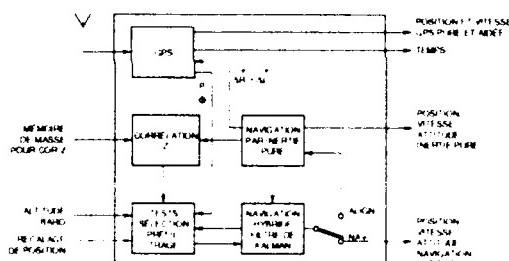


figure 6 - Bloc diagramme fonctionnel simplifié du logiciel de l'ULISS NG

Le couplage serré entre l'Inertie et le GPS comprend deux volets :

- l'aide de l'Inertie pure au récepteur GPS,
- le filtre de Kalman de navigation hybride qui reçoit des informations de différentes sources, dont le GPS intégré, mais également provenant de l'algorithme de corrélation d'altitude (TERCOR), du baro-altimètre ainsi que d'autres moyens de recalage (radar, systèmes infrarouges).

Ces informations sont d'abord sélectionnées et filtrées au moyen d'une fonction test/sélection dont le rôle est de détecter d'éventuelles dégradations ou pannes, rejeter les données non valides, sélectionner les meilleurs paramètres d'entrée et effectuer les préfiltrages nécessaires.

Le filtre de navigation hybride a deux modes de fonctionnement : le mode "Navigation normal" et le mode "Alignement" permettant un alignement rapide au sol, en v. ou à la mer.

Trois types de sortie sont disponibles :

- 1 - une sortie de navigation hybride, comprenant les positions vitesse et attitude issues du filtre de Kalman.

- 2 - une sortie Inertie pure (position, vitesse, attitude).

- 3 - une sortie GPS pur (position, vitesse)

Les sorties 2 et 3 sont des sorties secondaires.

De plus, le système délivre le signal 1 PPS (1 pulsé par seconde) ainsi qu'un temps GPS très précis

4.2 AIDE INERTIELLE SERREE AU GPS

L'aide inertuelle au récepteur GPS intégré comporte deux aspects :

- une aide à l'acquisition rapide à la mise en route du récepteur et une aide à la réacquisition en cas de masquage lors d'évolutions ou de brouillage.

- une aide en permanence aux boucles permettant d'améliorer la résistance à la dynamique et au brouillage du récepteur GPS.

- 1 - L'acquisition et la réacquisition rapide sont possibles grâce à l'initialisation des états des boucles de poursuite de phase de code et de phase porteuse avec des données suffisamment précises.

Ces données d'initialisation sont calculées à partir des position et vitesse du véhicule mesurées par l'inertie, des éphémérides satellite et du temps courant.

Pour chaque boucle de code C-A, on calcule les valeurs d'initialisation suivantes :

$$\rho(0) = \rho_a(0) \text{ modulo } 300 \text{ km}$$

$$\dot{\rho}(0) = \dot{\rho}_a(0)$$

Pour chaque boucle de phase porteuse, on calcule :

$$\Phi(0) = \phi_a(0) \text{ modulo } \lambda = 0$$

$$\dot{\Phi}(0) = \dot{\phi}_a(0)$$

où, pour chaque satellite poursuivi

- ρ et $\dot{\rho}$ représentent respectivement la pseudo distance mesurée et sa dérivée.
- Φ et $\dot{\Phi}$ représentent respectivement le Doppler intégré ou la mesure de phase de porteuse et sa dérivée.
- ρ_a et $\dot{\rho}_a$ représentent la distance antenne-satellite et sa dérivée calculées à partir des position et vitesse véhicule issues de l'inertie pure, des derniers éphémérides en mémoire et du temps

Le modèle simplifié des boucles de phase de code et de porteuse ainsi que les données d'initialisation (aide en acquisition/réacquisition), sont décrits en figure 7.

- 2 - L'amélioration de la résistance du récepteur GPS au brouillage et à la dynamique est réalisée grâce à une aide dynamique serrée des boucles de code et de porteuse

Cette aide en dynamique bénéficie de la bande passante importante et de la grande précision à court terme des mesures inertielles (force spécifique et angles d'attitude) pour calculer la composante "haute fréquence" de la dynamique de l'antenne GPS dans chacun des axes antenne-satellite. Cette

dynamique calculée est soustraite en temps réel dans les boucles.

Grâce à l'intégration physique du GPS très près de l'inertie et à l'interface directe rapide entre inertie et GPS, un contrôle serré est possible en temps réel, se concrétisant par une datation très précise, un retard minimal des informations et une aide inertuelle performante à haute fréquence. Cet avantage déterminant permet d'utiliser des boucles relativement simples du second ordre et conduit à une solution robuste et très performante

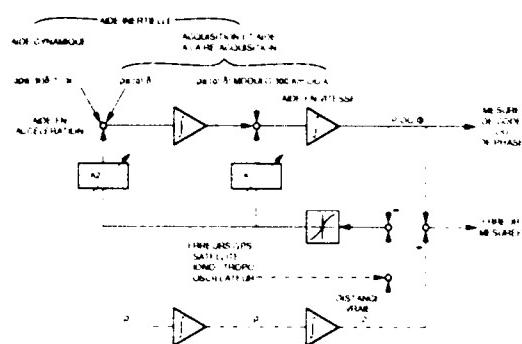


figure 7 - Modèle simplifié de l'aide inertuelle dans les boucles de code ou porteuse

Pour chacun des satellites poursuivis, on effectue le calcul suivant à 50 Hz :

$$\Delta \dot{\vec{p}}_a = \Delta \dot{\vec{p}}_a \frac{dt}{\Delta t}$$

$$\Delta \dot{\vec{p}}_a = \vec{u}^* \cdot \Delta \dot{\vec{p}}_a$$

$$\Delta \dot{\vec{p}}_a = \Delta \vec{R}_s \cdot (\Delta \vec{R} + \Delta \vec{L}) - (\Delta \vec{R} + \Delta \vec{L})$$

$$\Delta \vec{R} = \int_{t-\Delta t}^t (\vec{f} + \vec{g}) dt$$

$$\Delta \vec{L} = \Delta \vec{\omega}^{bi} \times \vec{L} + \vec{\omega}^{bi} \times (\Delta \vec{\theta} \times \vec{L})$$

où :

\vec{R}_s, \vec{R} sont les vecteurs position respectifs du satellite et de l'inertie,

\vec{u} est le vecteur unitaire de l'axe antenne-satellite,

\vec{L} est le bras de levier entre l'antenne GPS et le cœur inertiel,

\vec{f} est la force spécifique mesurée par l'inertie,

\vec{g} est l'accélération de la pesanteur.

$$\vec{\omega} = \frac{\Delta \vec{\theta}}{\Delta t}$$

est la vitesse angulaire de l'avion mesurée par l'inertie.

$$\Delta t, \Delta t$$

représentent les temps de cycle de l'aide inertuelle (20 ms) et du contrôle canal (1 ms).

$$\dot{\vec{x}} = \left(\frac{d \vec{x}}{dt} \right)$$

est la dérivée du vecteur \vec{x} dans l'espace inertiel [1].

Comme indiqué figure 7, l'aide en dynamique $\Delta \dot{\vec{p}}_a$ est une aide en accélération. Elle est réalisée à la fréquence de 1000 Hz. La précision, par rapport au temps réel, est meilleure que 1 ms, et quasiment aucune extrapolation n'est nécessaire. Les gains de boucle K1 et K2 sont adaptés en fonction de la dynamique mesurée, et ceci indépendamment pour chaque canal.

Là encore, l'intégration physique du GPS à l'inertie et le fonctionnement GPS "Tout en vue" en parallèle et synchrone sont très favorables.

4.3 NAVIGATION HYBRIDE SERREE

ULISS NG réalise une navigation hybride tridimensionnelle performante et sophistiquée grâce à un filtre de Kalman étendu unique.

Ce filtre est aussi utilisé pour l'alignement au sol, en vol ou à la mer, permettant un temps de réaction très rapide ainsi qu'une transition lissée et une initialisation précise du mode navigation. Il est basé sur une modélisation dynamique précise de la cinématique et des senseurs.

Le filtre de Kalman comprend les états suivants :

- position x, y, z,
- vitesse x, y, z,
- attitude x, y, z,
- biais accéléromètre en z,
- dérives gyromètres x, y, z,
- paramètres d'altitude barométrique,
- erreur d'horloge GPS et sa dérivée.

Le filtre de Kalman traite des observations multiples venant du GPS - mesures de pseudo-distance et phase de porteuse - du baro-altimètre, de la corrélation du terrain, d'autres moyens de recalage de position, ainsi que du recalage vitesse nulle en mode alignement. Le prétraitement effectué par la fonction test/sélection/préfiltrage est basé sur les indicateurs de validité et de qualité disponibles et sur les innovations et covariances issues du filtre.

Cette fonction réalise également un contrôle d'intégrité de l'inertie (détection de dégradation lente) ainsi que du GPS, pour le récepteur aussi bien que pour le segment espace.

Le filtre permet de calibrer les composantes jour à jour et à long terme des erreurs gyrométriques et accélérométriques. Il bénéficie de la résolution centimétrique de la phase de porteuse pour un alignement et une convergence rapides, ainsi que pour une vitesse hybride très précise.

Le filtre est conçu pour s'adapter à la précision et aux deux modes SPS et PPS du GPS (filtre bi-mode). Globalement, l'intégration fonctionnelle de l'inertie et du GPS - aide inertuelle et navigation hybride - est conçue pour garantir une très grande stabilité de la solution de navigation hybride, grâce à l'utilisation d'un algorithme

de Kalman-Bierman numériquement stable et à une corrélation long terme très faible des mesures utilisées, compte tenu de l'aide inertie pure directe et du filtre d'hybridation en boucle ouverte.

L'hybridation serrée INS/GPS fournit une navigation plus précise grâce à une modélisation fine des erreurs GPS et à l'utilisation optimale des informations GPS disponibles : il en résulte une meilleure efficacité des missions opérationnelles.

Les algorithmes d'hybridation ont été étudiés et mis au point sur simulateur. La figure 8 présente les performances obtenues en simulation avec une constellation de 21 satellites sur un scénario de vol typique d'avion d'armes construit à partir d'enregistrements de trajectoire réelle. Les courbes représentent les erreurs de position et vitesse encadrées de leur écart-type respectif ($+3\sigma$, -3σ) calculé par le filtre de Kalman.

5 - ESSAIS EN VOL SUR AVION DE COMBAT

5.1 PROGRAMME D'ESSAIS

Ce programme a été conçu en trois étapes, afin de remplir les objectifs suivants :

- démontrer en vol l'intégration physique Inertie-GPS et la performance du récepteur GPS code P(Y) à fréquence intégré dans l'inertie.
- démontrer ensuite, par ouverture du domaine de vol, l'apport de l'aide inertie serrée pour la disponibilité du récepteur GPS (et donc sa performance).

- enfin, montrer les bénéfices de la navigation hybride serrée : position et vitesse très précises, alignement rapide en vol et autocalibration des composants inertiel.

5.2 CONDITIONS ET REFERENCES POUR LES ESSAIS

L'ULISS NG est montée dans la baie équipement du Mirage III (voir figure 9).



figure 9 - Essais en vol sur Mirage III

La figure 10 montre le profil de vol type des essais sur Mirage III. Chaque vol a une durée d'environ une heure à une heure et demie, avec une vitesse moyenne d'environ 250 m.s. à une altitude de 15000 pieds. Les vols comprennent des phases de tonneaux enchaînés et de vol inverse, afin d'évaluer la vitesse de réacquisition du GPS après masquage. Des virages de 3 à 4 g ont également été effectués, en

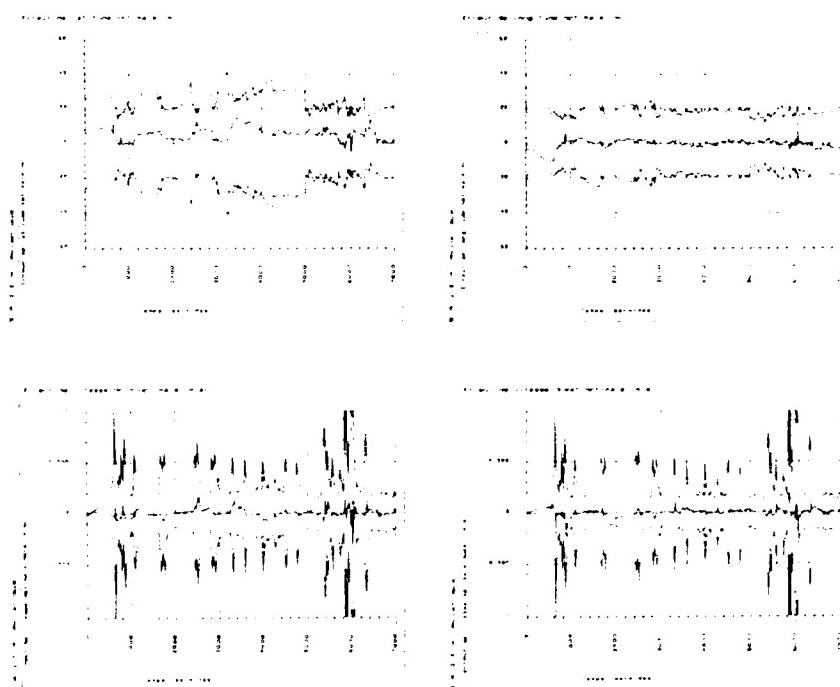


figure 8 - Erreurs de navigation optimale inertie-GPS couplage serré (position et vitesse)

restant en visibilité de certains satellites, afin de tester la tenue en dynamique du récepteur GPS intégré. L'instrumentation avion est composée d'une centrale ULISS 60 de grande précision, utilisée comme référence de navigation, d'un calculateur et d'un enregistreur de vol. La référence de position est donnée par plusieurs systèmes de trajectographie.

- un système de trajectographie laser, STRADA, dont la précision annoncée est de 1 m (1σ).
 - un radar de trajectographie.
 - une trajectographie différentiel dans un premier temps, et inertie/GPS différentiel dans un second temps.

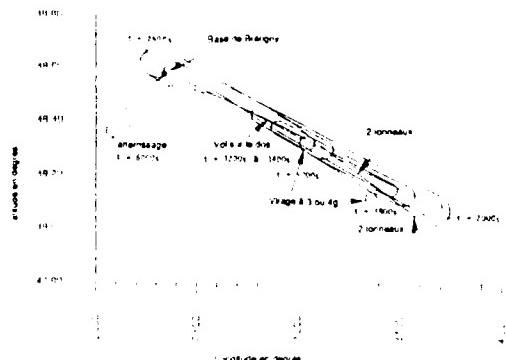


figure 10 - Profil de vol type

Les systèmes d'armes modernes devenant de plus en plus performants, leur qualification en vol nécessite un système de trajectographie très précis durant toutes les phases du vol d'essai, ce qui n'est pas toujours possible avec les systèmes de référence traditionnels.

La figure 11 présente le principe d'un système de trajectographie GPS différentiel composé d'un récepteur GPS multicanaux de référence au sol et d'un système Inertie GPS multicanaux à bord de l'aéronef. Une évaluation de la trajectographie GPS différentiel est en cours au Centre d'Essais en Vol de Brétigny. Les essais de trajectographie GPS différentiel en vol basse dynamique, commencés fin 1991, sur Caravelle au CEV de Brétigny, ont donné une erreur de position inférieure à 1 m à 1 σ (voir figure 12). Ces essais se poursuivent par une expérimentation en vol sur avion d'armes en haute dynamique.

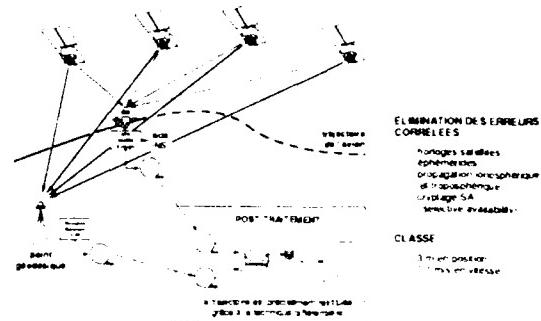


figure 11 - Trajectographie GPS différentiel
temps différé

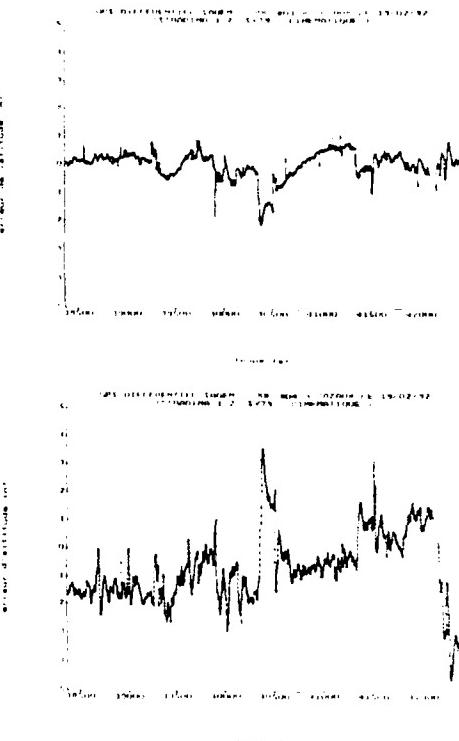
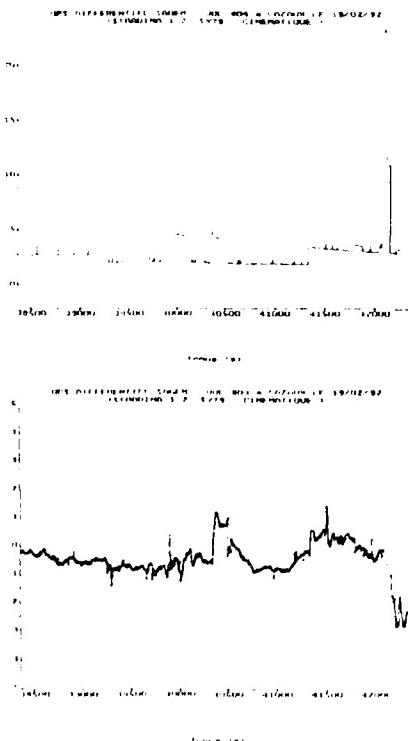


figure 12 . Trajectographie GPS différentiel Erreurs de position (m)

6 - CONCLUSION

Un récepteur GPS miniaturisé (0.7 l, 0.7 kg, 12 W) code P(Y) douze canaux a été intégré physiquement dans une centrale inertielles ULISS conduisant à un équipement compact léger "Tout en un" et "Tout en vue". fonctionnellement optimisé, en termes de performance et robustesse grâce à un couplage serré Inertie-GPS. Des vols d'essais sur avion d'armes sont en cours au Centre d'Essais en Vol de Brétigny. Grâce à la modularité de conception, une gamme complète de récepteurs GPS intégrés est dérivée de la version P(Y) L1/L2 par simple dépeuplement de l'électronique ou élimination de sous-modules du récepteur ; des versions P(Y) L1, C/A PPS, C/A SPS, six ou douze canaux sont ainsi disponibles. Ces nouvelles options GPS sont également intégrables dans les systèmes inertIELS nouvelle génération à gyrolasers de la famille SIGMA. Enfin, SAGEM a également développé un système de trajectographie à base de GPS différentiel, qui permet, de restituer la trajectoire de l'avion avec d'excellentes performances.

REMERCIEMENTS

Les auteurs souhaitent associer à cette publication, Mme D. DEMANGE, MM. L. GARIN, JP. GOLLIOT, S. GUICHARD, P. HOUBERT, JB. MINAZIO, P. QUICHAUD, B. ROGIC, R. TRAN qui ont été très impliqués dans les différents développements.

REFERENCES

- 1 Medium to High Accuracy Navigation with Pure and Hybrid Inertial Systems and Related Mission Planning, L. Camberlein and P. Lloret, GIFAS Conference 1989
- 2 Integration of GPS Receivers into Existing Inertial Navigation Systems, D A Tazarte and J G Mark, Navigation Vol. 35, 1988
- 3 Embedded GPS Solves the Installation Dilemma, M.A. Sturza and C.C. Richards PLANS 1988
- 4 An Option for Mechanizing Integrated GPS/INS Solutions, J A Soltz, J L. Donna and R L Greenspan, Navigation 1988
- 5 Development of a Tightly Integrated Ring Laser Gyro Based Navigation System C Kervin, R Chlossen, C Kiel and M Lynch, IEEE 1988.
- 6 GPS/Inertia: Navigation System Integration for Enhanced Navigation Performance and Robustness, R P Denaro and G J Geir, AGARD 1988
- 7 Inertie GPS : un mariage de raison à l'essai ., P. Lloret and B. Capit, AGARD 1986.
- 8 Benefits of Integrating GPS and Inertial Navigation Systems, T.N Upadhyay, B A. Kreigsmann and D B Cox Junior
- 9 ULISS G, a fully integrated "all-in-one" and "all-in-view" Inertia GPS unit, L. Camberlein, B. Capit - PLANS 90
- 10 Litton/Collins GPS Guidance Package, R E Ebner, T.J. Rickards, M.D. Yakos, W A Mickelson - PLANS 92.
- 11 The Texas Instruments Honeywell GPS Guidance Package, P W Ward, Dr. Mahesh Jurage - PLANS 92.
- 12 Architectures and GPS INS Integration Impact on mission accomplishment, Z H "Stan" Lewantowicz, WP AFb - PLANS 92

RETROFITTING OF GPS INTO EXISTING NAVIGATION SUITES

D.I. Callender, N.F. Watson
GEC Ferranti Defence Systems Ltd
Navigation Systems Division
Silverknowes
Edinburgh EH4 4AD
Scotland

SUMMARY

As GPS signal availability reaches operationally useful levels, and in particular following experience of its usefulness in the Gulf operations, widespread requirements are beginning to arise for the incorporation of GPS into existing combat aircraft.

A considerable amount of study and development work has been carried out by GEC Ferranti to investigate different approaches to integrating GPS into existing navigation systems with the minimum impact on installation, interfacing and operating procedures.

This paper describes some different approaches to integrating GPS together with their relative merits and deficiencies. Two practical systems are described in detail and some simulation and trials results are presented together with some aspects of the GPS integration which will form the basis of future development work.

1 SYSTEM ARCHITECTURE

The GPS receiver can be linked to existing avionics in a variety of ways depending on the type of systems with which it must interface. This can range from a stand alone receiver installation, with no interface to any of the existing avionics, to a fully integrated fit with the receiver embedded in another unit in the avionics suite.

1.1 SIMPLE NAVIGATION AID

GPS can be used as a simple navigation aid when the existing navigation fit consists of old equipment with analogue interfaces or a fit with a relatively simple main computer or FMS. This is not an optimal fit but may be the only method of fitting GPS to older aircraft.

1.1.1 Stand Alone Receiver

This is the simplest fit (at least as far as interfacing is concerned) with the receiver connected to a CDU, an antenna and a source of power (Figure 1.1). It is used as just another source of position and possible steering information which is used manually by the pilot or navigator. However, GPS is not used to its full advantage because it is not

interfaced with any of the existing avionics. This leads to problems during aircraft manoeuvres when the satellites are masked from the antenna and the receiver can no longer provide a navigation solution. When the aircraft completes the manoeuvre and is flying level again, the GPS receiver may be unable to re-acquire the satellite signals because the Doppler shift of the carrier due to vehicle motion may be outside the tracking loop acquisition capability. The operator then has to manually insert course and speed information into the GPS via the CDU to allow it to re-acquire the satellite signals. This is a time consuming business if it happens frequently and can result in the GPS becoming useless during high dynamic flight.

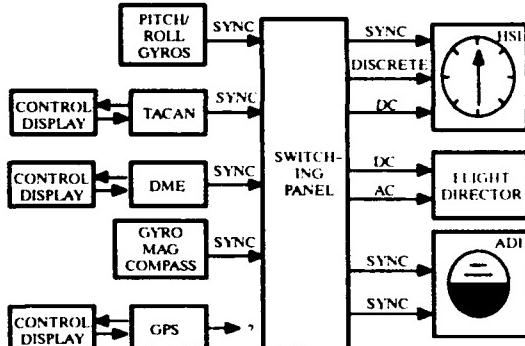


FIGURE 1.1 STAND ALONE GPS FIT

This type of fit was used on the Tornado aircraft during the Gulf War. The data was used by manually inserting the GPS position into the main computer as a fix and thus updating the navigation system. There is considerable scope for inserting an incorrect fix, especially when under pressure. In addition, the GPS suffered from the re-acquisition problem described above during manoeuvres that resulted in a high change of velocity vector.

This type of fit should only be considered for a very benign environment or as a policy of desperation.

1.1.2 Input To A Main Computer

The GPS can be integrated via a main computer or FMS in more modern avionics fits which have some level of digital interfacing although analogue

interfaces may still be used for the flight instruments and flight control systems. The digital interfaces are usually implemented as point-to-point links between the FMS and the avionics (Figure 1.2).

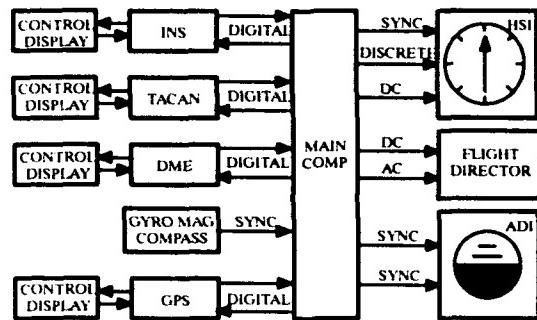


FIGURE 1.2 GPS INTEGRATED VIA MAIN COMPUTER

The computer or FMS must have a suitable spare interface and the necessary software to make use of the GPS data or the GPS must be configured to mimic an existing navigation aid which it then replaces. This is an improvement over the stand alone fit but it still does not make best use of the GPS data. If no aiding information is provided from the FMS to the GPS then this configuration will suffer from exactly the same problems during manoeuvres as the stand alone system. However, if aiding is provided by the FMS, then there is a significant gain in that the operator no longer has to manually insert data before the GPS will re-acquire satellites. This allows very rapid re-acquisition within 1 or 2 seconds after the satellites are visible compared with an indefinite delay for an unaided GPS.

Most existing FMS or computer based systems are not configured to do anything with the GPS data other than use it as an alternative source of position data for display or steering calculations. This data will only be available during level or low dynamic flight while 4 or more satellites are visible.

1.1.3 Fixing Aid

A more capable FMS or computer based avionics suite is likely to have the ability to use the GPS data in the navigation solution together with other sensors and to aid the GPS with data from the navigation solution. The GPS can also be used as an alternate source of prime navigation data.

This makes better use of the GPS data and overcomes the disadvantages of the simpler schemes. However, the FMS software will generally be written to be able to use a variety of GPS receivers and other sensors and therefore the performance available from integration with this type of configuration may not be the best that can be achieved.

This type of system is likely to be bus orientated within the navigation suite while keeping point-to-point interfaces (both analogue and digital) with the aircraft instruments and flight control systems (Figure 1.3). This makes the installation of the GPS relatively easy since it can be linked to the rest of the system by coupling it to the 1553 data bus. Control and display functions are carried out by the FMS through CDUs on the 1553 bus.

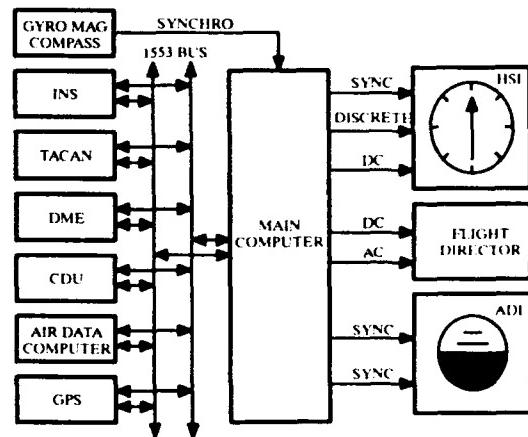


FIGURE 1.3 GPS INTEGRATED VIA A 1553 BUS BASED SYSTEM

Introducing GPS to this type of navigation suite gives a significant improvement in the navigation performance for a moderate cost, provided the FMS is already capable of making use of the GPS data on the 1553 bus or other available interface.

1.1.4 Input To Mission Computer And MFDs

The latest navigation suites are based on a 1553 bus orientated system with one or more mission computers and a number of multifunction displays (MFDs) (Figure 1.4). In principle, it is easy to couple the GPS to the 1553 bus, but . . . the mission computer may not have been programmed for the GPS bus traffic or integration of the GPS data and also may not have the necessary software to provide control and display facilities for the GPS via the MFDs. Providing this software can be a very expensive and time consuming business. Under

these circumstances, alternative methods of incorporating the GPS should be investigated.

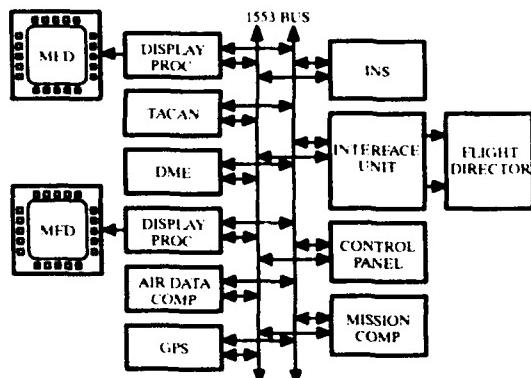


FIGURE 1.4 GPS INTEGRATED WITH DIGITAL AVIONICS

1.2 INTEGRATION

The way in which the GPS is linked to the existing navigation suite is only half the story. Best use should be made of this highly accurate (but not necessarily continuous) source of navigation data. In addition, reasonable measures should be taken to maximise the time for which GPS data is available.

The GPS data can best be used by integrating it with information from one or more other sensors. The most common sensor used for this purpose is an Inertial Navigation System (INS) and the integration is usually carried out by a Kalman Filter.

1.2.1 INS Error Model

The Kalman filter contains a mathematical model of the INS error sources which describes their effect on the navigation outputs from the INS. There are generally a large number of error sources that could be included in the filter, but not all of them are necessarily observable or have a significant effect on the INS performance given typical flight profiles. Further, the number of Kalman filter error states impacts on the computation time required for a filter cycle and also the memory requirements.

The error states to be incorporated in the filter can be identified through knowledge of the INS and by carrying out a covariance analysis using a truth model driven by typical flight profiles. The truth model consists of a Kalman filter with all significant INS error sources included.

GPS error characteristics must also be considered in the Kalman filter design.

1.2.2 INS Calibration

A by-product of the integration using a Kalman filter is an estimate of some of the INS instrument errors. These can be held after a flight and applied to the INS on the next flight. The covariances from the filter also give a measure of the accuracy of the estimates and they can therefore be used to provide a weighting factor for the stored update.

The calibration is likely to be for gyro drifts or accelerometer biases and scale factors but may also include gyro mass unbalances if the integration is sophisticated enough.

Automatic calibration of the INS may give a significant reduction in the cost of ownership of the INS, particularly older gimballed systems. The INS will always be well calibrated and will therefore not require scheduled maintenance and is also less likely to be removed from the aircraft due to poor performance.

1.2.3 Aiding The GPS

The GPS requires velocity or course and speed data to aid re-acquisition of the satellites after loss of the signal due to a manoeuvre or terrain masking. It also provides an improvement in the anti-jamming capability of the receiver because the tracking loop bandwidth can be reduced, thus rejecting some of the jamming signal.

The GPS velocity is measured at the antenna position. The aiding data is usually referenced to a different location and therefore the data must be corrected for lever arm effects using attitude rate data.

Attitude rates tend to be rather noisy and this can result in significant noise being introduced to the otherwise quiet velocity data. The aiding data may therefore need to be filtered, but this in turn introduces delays to the aiding data which may not be acceptable.

The characteristics of the aiding data also have to be considered together with the way in which the GPS receiver uses it. For example, is the GPS sensitive to the absolute accuracy of the aiding data or does it calibrate out offsets? What about the effect of delays in the data?

The best source of velocity data is likely to come from the INS. Errors in the velocity generally change slowly with time, allowing the GPS receiver to track them (if it is so designed). The data also has a high bandwidth, allowing it to track vehicle motion with little delay, and is continuous (unless the INS fails!).

1.2.4 Location Of The Integration Process

There are several possible locations in which the integration can be carried out, especially in the more modern avionics suites which have multiple processors with significant spare capacity. Several of these will now be considered.

1.2.4.1 Main Computer

This may be the most obvious option, particularly if the software is already written for it. However, the integration filter in pre-written software may not be optimised for the specific INS and GPS being considered for the system.

The reversionary characteristics of such a system are not good in the event of a main computer failure. The navigation solution may revert rapidly to raw INS and the corrections for the INS will be lost unless they have been stored elsewhere periodically.

1.2.4.2 GPS

The GPS receiver may also be an appropriate location for the integration process. Many receivers already have basic Kalman filters designed to estimate INS errors on the rate aiding data. These filters tend to be limited to a few error states and are generic in nature and are therefore not well suited to the INS calibration aspect of the integration. The reversionary capability in the event of GPS failure is no better than that for the main computer.

1.2.4.3 INS

The last option to be considered uses the INS to perform the integration function. This has a number of significant advantages. The first is that the Kalman filter will be optimised for the type of INS in which it is embedded, thus obtaining good calibration of the INS errors. Secondly, attitude data (and possibly attitude rate data) is readily available for lever arm correction of the rate aiding data. In addition, the rate aiding output can be specifically designed to provide low latency data to the GPS by careful time synchronisation of the data processing with other activities inside the INS. The same also applies to the time synchronisation of the incoming GPS data, which can be done more accurately when there is total control over the timing of events in the INS.

Lastly, the reversionary characteristics are excellent since the system is unaffected by failure of other aircraft systems provided that the interface between the INS and GPS remains operational. If the GPS fails, the INS outputs are still corrected by the propagated error states from the Kalman filter.

If the INS fails, the Kalman filter is not required anyway.

1.3 CLOSE INTEGRATION

There are significant benefits which can be gained by having the GPS embedded within another unit which may already be part of an avionics system. This will reduce installation costs and also allows a more sophisticated integration scheme to be used. On the down side, the cost of development of such a system is significant.

1.3.1 Single Box Solution

The incorporation of the GPS receiver within an existing unit greatly reduces the installation costs of the GPS per aircraft. Indeed, this may be the only method of fitting GPS to aircraft where space and weight are at a premium.

The GPS can be embedded as a separate unit, operating from its own power supplies and maybe also having its own interfaces to the rest of the avionics for reversionary facilities. However, it will be linked to the INS via a dedicated internal interface, for example dual port RAM shared by the GPS and the INS. This allows rapid transfer of large quantities of data and good control of the time synchronisation processes.

Alternatively, the GPS can be incorporated as a card set, using the INS power supplies and interfaces. The reversionary capabilities are likely to be poorer than in the previous case since there are several common mode failure mechanisms, but it will be cheaper, lighter and consume less power.

1.3.2 Use Of Pseudo Range And Range Rate Data

The single box solution also allows for the possibility of using pseudo range and range rate data because of the internal interface. This is really of benefit for military systems using PPS GPS data because of the security implications of using this data once the effects of SA have been removed. The internal interface means that no sensitive data is transmitted outside the unit thus avoiding the requirement for Tempest clearance of the aircraft.

The benefits of using pseudo range and range rate are likely to be most evident for high dynamic and vehicle applications, where continuous visibility of four satellites may not be possible. The GPS will only give a navigation solution for a small proportion of the time under these circumstances. However, useful information is still available even when only one satellite is visible and the Kalman filter can be designed to make use of individual satellite ranges and range rates instead of the GPS navigation solution.

2 CASE STUDY 1 – FIN 1075G

The first case study is of an INS/GPS integrated system designed for use on the Harrier GR7s of the Royal Air Force. The INS that was already fitted was a GEC Ferranti FIN 1075 system which was capable of modification to become part of an integrated system.

2.1 INTEGRATION APPROACH

The avionics suite consists of a mission computer which communicates with other systems over a dual redundant 1553 bus. All control and display functions are handled by the HUD, two MFDs and an up-front control panel (Figure 2.1).

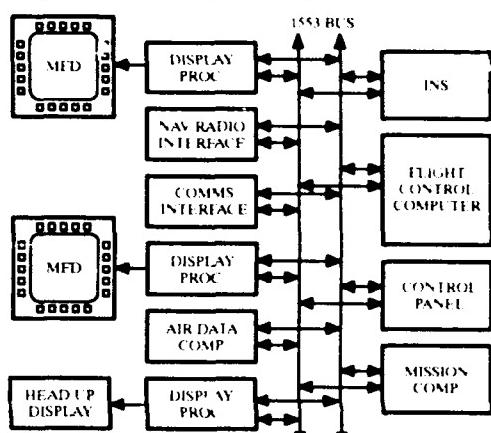


FIGURE 2.1 HARRIER GR7 AVIONICS ARCHITECTURE

2.1.1 Factors Considered

The first factor to be considered was the way the GPS would interface with the rest of the system. The 1553 interface would seem to be the obvious choice, but the time and expense of writing an update to the mission computer software (which is the bus controller) precluded this approach. This also eliminated any thought of performing the integration in the mission computer.

The GPS therefore had to interface directly with the INS since the 1553 data could not be modified. In addition, the INS would be required to perform the integration. A 68020 card with the necessary processing capability for the integration had already been designed for another project. It also had two RS422 ports on it which would allow direct communication with the GPS via the instrumentation port. Thus a single additional card in the INS would carry out the integration and all the GPS interfacing.

The Harrier GR7 has an integrated cockpit with the majority of the avionic systems being controlled through the Up Front Control Panel and the

MFDs. There is very little space available to allow the fitting of dedicated CDU for the GPS and this also goes against the design philosophy of the cockpit. The only dedicated control that could be provided was a power switch, a GPS Fail indicator and a GPS Ready indicator. Any other information would have to be made available to the pilot using existing facilities in the INS.

The INS generates several maintenance pages which drive the MFDs directly. One of these pages was modified to display GPS data instead of INS data. This display was only available on the ground – or so it was thought. The initial flight trial demonstrated that this page could be displayed in flight provided that it was selected before take-off and that no other display was selected on the particular MFD during the flight. There was thus some GPS information available on the pilot during flight which was useful for the interpretation of the HUD video of the sortie.

The system also provided a data logging output on an RS232 interface. A hand-held PC was used to record this data during flight for subsequent processing and analysis.

2.1.2 Architecture

The integrated system consisted of a separate GPS receiver linked to the INS by a bidirectional RS422 interface. The integration was performed in the INS on the additional processor card. In order to avoid changes to the 1553 data format, the INS navigation parameters were modified by the output of the Kalman filter before they were transmitted over the data bus. The rest of the avionics therefore received data as if from a very good INS (with some rather nice INS like characteristics).

The Kalman filter was loosely coupled to the INS in the sense that the error estimates were not fed back to the INS (Figure 2.2). This was done to avoid possible stability problems together with the concern of the customer that the system performance should never be worse than that for a standard INS. By keeping the correction open loop, the raw INS data was always available as an independent back-up navigation source.

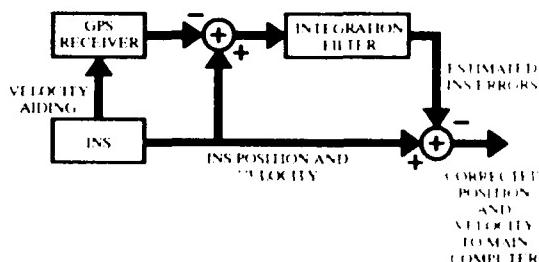


FIGURE 2.2 INTEGRATED INS / GPS SOLUTION

State estimates from the Kalman filter were used to correct the errors in the horizontal channels only. The attitudes were not corrected for tilt errors because of flight safety considerations, since the INS is the primary source of attitude information on the aircraft. Similarly the height channel was not considered for integration because of flight safety considerations.

The time scale for the overall project was very short (6 months) and therefore a low risk solution had to be considered. This meant restricting the number of INS error states and using GPS position and velocity data rather than investigating the use of pseudo range and range rate. It was recognised at the start that this might lead to lack of GPS availability during manoeuvring or low level flight, but this was considered an acceptable limitation for the first phase of the programme.

The integrated system as finally configured required very little in the way of modification of the INS. The additional processor card acted as a co-processor to the existing INS processor, thus keeping the software associated with the integration largely separate from the INS navigation software (Figure 2.3).

2.1.3 Installation

The system was designed to have as little impact on the aircraft as possible. Obviously, additional wiring and mounting fixtures were required to accommodate the GPS receiver and the antenna. The GPS receiver also had to be linked to the INS, which required the removal of the wing and engine of the GR7. This accounted for approximately 70% of the total installation cost.

The antenna was mounted on the upper surface of the fuselage at the trailing edge of the wing. This site was used because it was convenient and so may not be the best position on the aircraft. The field of view of the antenna is reasonable except looking forward where the cockpit and upper fuselage give some obscuration of low elevation satellites.

The GPS receiver and the data logger were mounted in the rear equipment bay and were accessible to the ground crew via the Daily Access Panel inside the main wheel well. This is obviously a trifle inconvenient when loading crypto keys or retrieving the data logger, especially on a wet day! This aspect will be addressed during any further modifications to the installation.

The INS was a form and fit replacement for the existing INS with the exception that the test connector was used to output data to the data logger.

Lastly, the pilot was provided with a control panel in the cockpit consisting of a power switch and two indicator lights. The first light indicated a GPS BIT failure while the second one indicated that the GPS was providing a navigation solution with a radial error of less than 100 metres (FOM less than 3).

The pilot display consisted of one page of data accessible from the maintenance menu. This allowed visibility of the GPS position, height, time, velocity, number of satellites used in the navigation solution and the figure of merit (FOM) (Figure 2.4). In theory, this page was only accessible on the ground but it was found that if it was selected prior to take-off, it could be displayed throughout the flight provided no other page was selected on the MFD.

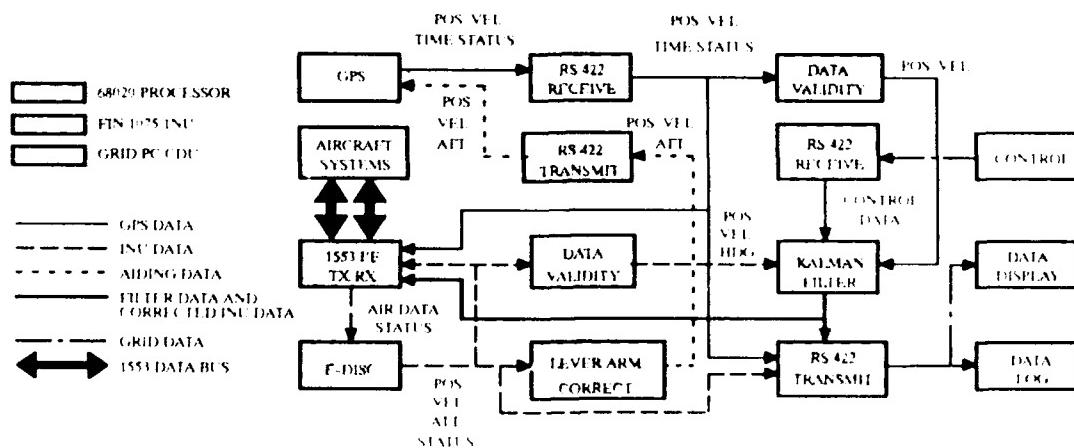


FIGURE 2.3 INS/GPS DEMONSTRATOR DATA FLOW DIAGRAM

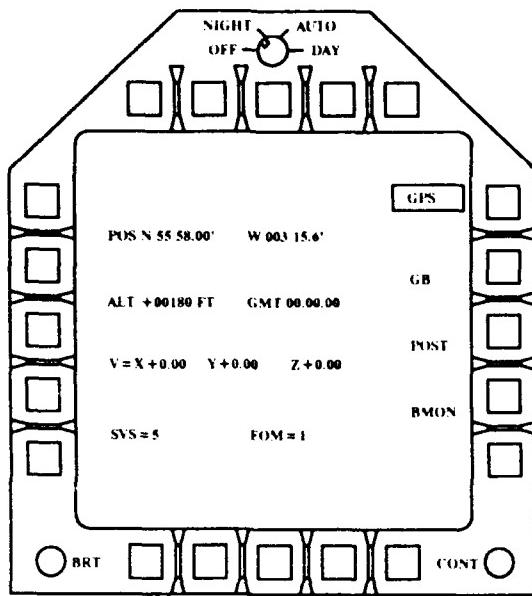


FIGURE 2.4 MFD GPS DISPLAY PAGE

The integrated system was also flown on GEC Ferranti HS125 aircraft and a Hercules C130. Neither of these aircraft were equipped with a 1553 based avionics system with the necessary MFDs to simulate the Harrier GR7. An additional RS422 output was therefore provided from the INS to drive a CDU software package on a laptop PC. This CDU program allowed the system to be controlled to explore various operating aspects such as the use of rate aiding by the GPS and the frequency of the Kalman filter measurement updates.

2.2 INTEGRATION FILTER DESIGN

2.2.1 The Requirement

The FIN1075G programme was required to provide a short-timescale, low-risk integration of GPS into the GR7 with minimal modifications to existing standard equipments or operating procedures. The ultimate in integrated Navigation position accuracy was not required - with a relatively modest 0.2 nautical mile worst-case error limit being adopted as the design aim. It was assumed that periods of non-availability of at least four GPS satellites would result mainly from airframe masking rather than terrain masking and would be relatively short and infrequent. The design aim was to meet the required position performance during absence of a full GPS solution for periods not exceeding 10 minutes.

For a number of reasons - including availability within the timescales of this programme - a GEC

Plessey Avionics GPS receiver was selected. The characteristics of this GPS receiver had to be considered in optimising the filter. The receiver operation is described below.

2.2.2 The GPS Sub-System Operation

2.2.2.1 Operation With Full GPS Solution

When at least four satellites are being tracked with good GDOP and signal-to-noise ratio this receiver operates essentially as a stand-alone navigator.

Pseudo-range and Pseudo range-rate measurements are used to update estimates of vehicle position and velocity and receiver clock error states in a Kalman Filter (KF). This KF is augmented to include acceleration states in a "high dynamic" vehicle such as GR7 - giving 11 states in all.

Although it operates entirely "stand-alone" in this mode the GPS receiver is designed to make use of velocity aiding data from an INS when available. The velocities are first corrected for lever arm effects by the INS so that the aiding data provided to the GPS relates to antenna motion.

The relatively low bandwidth errors in the INS velocity aiding data are modelled in a second Kalman Filter within the GPS. The inputs to this "INS Calibration" filter - INS velocity error measurements - are formed by subtracting GPS velocity estimates from the INS aiding data. The filter estimates the parameters of a simple INS error model. This data is of value during any subsequent periods when the GPS requires aiding.

In addition any barometric altitude aiding data is compared against the GPS height solution to determine a bias correction. This data is also of value during any subsequent periods when the GPS requires aiding.

2.2.2.2 Operation Without Full GPS Solution

Under conditions in which insufficient high quality satellite data is available to form an independent GPS solution due to satellite masking or poor GDOP, the GPS receiver continues to track as many of the available satellites as possible.

If no aiding data is available the Navigation Kalman filter is modified to estimate only five states (3 position and 2 velocity states) using all available satellite pseudo-range and pseudo range-rate measurements and bias corrected barometric altitude data.

Lack of aiding data can lead to difficulties in reacquiring satellites following manoeuvres during which little satellite data is available. There may be insufficient data to keep track of velocity changes during the manoeuvre. Large velocity changes may

make satellite reacquisition slow or even impossible when satellites again become visible. In this case a crude estimate of aircraft speed and heading inserted manually via a CDU may be necessary to allow reacquisition of satellite tracking.

When aiding data is available, the incoming INS velocity data is first corrected by the INS errors estimated and propagated in the INS calibration Kalman Filter within the GPS receiver. It is then used in the GPS receiver to aid the tracking loops and the navigation solution.

Velocity aiding data from the INS, further compensated for lever-arm effects, is used in a 'feedforward' sense to aid the carrier tracking loops, thus allowing lower loop bandwidth to be selected improving signal to noise ratio and anti-jam capability. As a result the GPS velocity outputs are correlated with INS velocity at frequencies above the (reduced) tracking loop bandwidth.

When the navigation solution is being aided, satellite pseudo-range rate data is no longer used to update the GPS Navigation Kalman Filter. Corrected INS velocity aiding data is integrated to give a position estimate which is updated by pseudo-range measurements from any available satellites. Thus when INS aiding data is being used the errors in GPS output data become correlated with the errors in INS aiding data - with a potentially long correlation time.

2.2.3 The Inertial Sub-system

The FIN1075 inertial system is a gimballed INS employing three GEC Ferranti Type 125 floated Rate Integrating Gyros and three pendulous accelerometers in a thermally controlled environment. The instrument axes are maintained nominally in a locally level orientation displaced about the vertical from North by a Wander angle which rotates at the vertical transport rate.

A conventional Baro-Inertial Height Loop provides vertical axis navigation data.

This product is a derivative of earlier GEC Ferranti INS systems with a long service history and as a result the error model of the INS is well understood. Dominant terms (for most short flights) are the three gyro drifts and the resulting tilts and velocity errors - including those tilts resulting from the levelling and gyrocompassing phase.

As a medium accuracy system (in the 1nmph class) tilt errors are relatively small following a full alignment and the INS error propagation is well approximated by a linear error model.

2.2.4 Integration Filter Mechanisation

As explained elsewhere in this paper, for ease of modification it was decided to locate a Kalman Filter within the FIN1075G INS incorporating GPS Position and Velocity data with INS measurements to produce combined INS/GPS output data in place of the conventional INS output data. (See Figure 2.2)

INS velocity measurements are also provided to the GPS receiver to improve satellite reacquisition and tracking loop performance.

2.2.5 Baseline Integration Filter Design

"Velocity Error" and "Position Error" measurements are formed by differencing Position and Velocity estimates from the two navigation sensors. The GPS position and velocity data are available to the full PPS precision. The use of S/A corrected pseudo-range and pseudo-range rate data is avoided for security reasons.

More fundamental errors contributing to the measurement differences (such as gyro drifts, tilts etc.) are estimated by a Kalman Filter incorporating an error model of the INS as described below.

The "best estimates" of INS velocity and position error are propagated and subtracted from the INS output to provide wide-bandwidth, continuous data with long-term accuracy similar to GPS.

Estimates of INS errors are available for updating the INS calibration stores post-sortie, however in-flight feedback of INS calibration data is avoided to maintain the integrity and independence of the INS attitude reference.

2.2.5.1 GPS Measurement Model

Within the Integration Filter GPS Position and Velocity outputs are considered independent, and their errors are treated as "white noise" - since GPS bias and other correlated errors are both difficult to observe and sufficiently small to ignore using the PPS measurements from the PA9052.

2.2.5.2 Measurement Update Rate

It was considered desirable only to use GPS data which is totally independent of INS aiding data to avoid "cascaded filter" considerations and potential problems with filter divergence due to possibly unstable feedback loops. The required operational scenarios are such that periods of non-availability of a four-satellite constellation exceeding a few minutes are expected to be rare. Also the error growth-rate of an open-loop FIN1075 INS would be expected to be of the order of 200m (worst-case) in the first 15 minutes immediately following alignment. Therefore integrated system performance was expected to meet the 0.2 nm

requirement throughout periods of non-availability of GPS aiding data of the order of 15mins given that INS errors remain modelled in flight to the level achieved following a ground-based alignment phase.

FIN1075 errors are dominated by low-frequency effects which remain consistent over relatively long periods of time. It was felt that such errors may be estimated adequately using GPS data relatively infrequently.

Therefore the baseline assumption was made that GPS aiding data would be used to update the Kalman Filter relatively infrequently - perhaps every 200 to 300 seconds - and only at times when a pure GPS Navigation solution was available.

Several advantages were expected to be gained by the decision to use GPS data relatively infrequently:-

- (1) Only well validated GPS measurements would be used following a period of successful completion of error and data validity checks - avoiding the inclusion of "rogue" measurements in the integration filter.
- (2) The assumption of "whiteness" in the GPS measurement data is more valid if the time between measurements is long compared with the true correlation time - a function of tracking loop bandwidth, GPS Navigation Kalman filter dynamics etc. (NB This may be more important when using SPS data).
- (3) Several consecutive error measurements may be averaged to reduce the true white noise component of error measurements introduced by the GPS receiver, the INS and any synchronisation jitter.
- (4) The simple GPS white-noise error model would be more tolerant to gaps in GPS availability of several minutes if the GPS measurement noise covariance level was selected on the basis of a low update rate.

However following a series of covariance simulations intended to determine preferred values for filter noise parameters, states and update rates, it was shown that more frequent updates gave lower "step changes" in integrated system position and Velocity outputs .

This is believed to result from an improvement in the observability and separation of different INS

error mechanisms with more frequent measurements. An update rate of 1/30 Hz was finally chosen.

2.2.5.3 INS Error Model

As a result of a series of covariance simulations using a comprehensive GPS and INS truth model with the baseline Kalman Filter the following ten Kalman Filter States were chosen:-

Horizontal Position Error	(2 states)
Horizontal Velocity Error	(2 states)
Horizontal Tilts	(2 states)
Heading Error	(1 states)
X,Y,Z gyro drifts	(3 states)

It was decided not to include accelerometer bias, and misalignment terms since these errors remain fairly well correlated in most flight scenarios (where there is only a small change in wander angle). The absence of these terms is neither significant in determining integrated navigation performance nor in maintaining calibration of the Inertial System.

2.2.5.4 GPS Data Synchronisation And Validation

Formation of accurate position and velocity error measurements is critical for proper operation of the Kalman Filter. Care was taken to ensure that data from the two independent sensors was adequately synchronised, and that only valid measurements from the GPS were used.

GPS measurements are made at 1Hz and the subsequent GPS position and velocity estimates transmitted on the instrumentation port up to 400mSec later. The INS position and velocity estimates are formed from measurements taken at regular intervals occurring at approximately 40Hz. The INS position and velocity outputs pertaining to the two measurement points prior to and subsequent to the GPS measurement are recorded, and linear interpolation is used to generate an estimate of the INS measurements appropriate to the time of the GPS measurement. Thus error measurements can be formed from effectively synchronous data, becoming available up to 400mSec after their true time of validity.

Prior to use by the Kalman Filter a number of data validation checks are performed. These involve monitoring GPS mode, "Health" and Covariance estimate outputs, and comparison of the individual proposed Kalman Filter innovations against a limit determined from the predicted innovation covariance.

2.3 RESULTS

2.3.1 Vehicle And HS125 Trials

The system was initially tested in a vehicle driving around Edinburgh. This gave a feel for the ability of the rate aiding to assist the GPS in recovering satellite tracking after obscuration by buildings. This was the first time that the GPS receiver had used rate aiding data over the RS422 instrumentation port. Some minor problems were found but on the whole the system performed very much as expected.

The system was then fitted in the GEC Ferranti HS125 for a series of flight trials. The first flights were conducted using the GPS in a stand alone mode to try to assess the effects of satellite masking during manoeuvres and the benefits of using rate aiding data. The flight profiles consisted of patterns giving high changes in velocity vector in one or both axes which were conducted at low and high bank angles with and without rate aiding.

The results indicated that obscuration was a significant problem especially when rate aiding was not used. Under these circumstances, the receiver could take several minutes to re-acquire the satellites once the aircraft was flying straight and level. When rate aiding was applied, the GPS receiver tended to have all five satellites back in track before the aircraft had completed the roll out from the manoeuvre.

A total of twelve hours flying was carried out during these initial tests which were used to debug the system software and check the data recording system before the equipment was delivered to A&AEE Boscombe Down for flight trials on the Harrier GR7 and C130 Hercules aircraft.

2.3.2 Harrier GR7 Trials

A Harrier GR7 of SAOEU at A&AEE Boscombe Down had been modified for the flight trials of the INS/GPS system. These trials were intended to establish a baseline standard against which future flights could be assessed and then go on to explore the behaviour of the INS/GPS system in an operational environment, including weapon release trials and consideration of tactical benefits.

The first flight in the GR7 was conducted on the 19th December 1991 with moderate quality GPS coverage. The flight lasted for one hour and was (for a first flight) very successful. The Kalman filter had only used the GPS data for the first thirty minutes before the GPS Estimated Position Errors exceeded the rejection level and the GPS data was no longer used by the filter. Up until that point, the

errors observed on flying over waypoints were zero to within the steering calculation resolution.

Some minor problems were observed during the first flight with the GPS receiver and the INS software. Considerable effort was put in by Ferranti and Plessey to resolve these problems before flying started in earnest in January 1992.

A total of eleven flights were flown for the initial evaluation of the system. Data from these flights was recorded by the data logging system (until the batteries went flat) and also on the HUD video recording system which included the pilot radio/intercom. The pilot could therefore make comments about the quality of GPS data just prior to overflying the waypoint to help with subsequent analysis. The accuracy of the system was estimated by measuring the offset of the target marker from the waypoint on the HUD video.

There were several occasions when the system was observed to be in error by up to 0.3 of a mile on overflying a waypoint. This was largely due to an incorrect position having been quoted for the waypoint. Thus a significant fraction of the observed errors in flight may be due to the resolution of the waypoint data and the steering calculation.

Typically, the overflight errors were estimated to be of the order of 100 metres with terminal errors on the pan of 30 to 60 metres. These initial flights were flown without the crypto variables having been loaded into the receiver because of operational limitations. This will be addressed in time for the next phase of the trials.

Finally, some mention should be made of the weapon release trials conducted during a night flying exercise.

The INS/GPS equipped GR7 flew low level at night to a bombing range in the north of Scotland and released a live bomb, scoring a direct hit. This was the first demonstration of the GR7 ability to achieve the night attack role with INS/GPS fitted.

2.3.3 C130 Hercules Trials

In parallel with the GR7 trials, a second system was flown in a C130 to provide a quantitative assessment of the integration. This time, the receiver was loaded with the crypto variables and was therefore providing PPS data. A reference system was also flown against which the INS/GPS could be compared. These flights also explored the effects of signal attenuation and masking on the GPS receiver and the integration process.

The data from these flights is still being processed, but some factors have already emerged. The

Kalman filter has not yet been "fine tuned" and therefore may not be settling as rapidly as it should. There is also some doubt as to the way in which the GPS receiver behaves when using rate aiding data. When the analysis of the flight data has been completed, the filter parameters will be modified and a limited flying programme conducted to verify that the changes behave as expected.

3 FUTURE WORK

The integration work carried out so far is just the first step in producing an operational INS/GPS integrated system. It is a basic solution which can be improved considerably both in terms of performance and of the capabilities that it can provide.

3.1 USE OF PSEUDO RANGE AND RANGE RATE DATA

The integrated system as it stands makes use of GPS position and velocity data. In order to provide this data, the GPS receiver must track four or more satellites. If the aircraft spends much of its time manoeuvring or flying at low level with significant masking terrain, the GPS will be unable to provide a navigation solution for much of the time. However, it is still receiving useful information from the satellites that are still visible. The Kalman filter can be modified to make use of individual satellite range and range rate measurements instead of the position and velocity data.

There are security problems associated with the transmission of the data necessary for the Kalman filter to make use of the range and range rate over a data link. The security aspect has still not been resolved and may present an insurmountable problem to carrying out this work with this configuration of system. An embedded system would provide the solution to this problem since the sensitive data does not leave the unit.

3.2 FEEDBACK OF ERROR STATES

The present configuration applies the error states to the INS data before it is transmitted to the rest of the aircraft systems. The error states are not fed back to the INS and do not affect the raw inertial behaviour of the INS. The system performance can benefit from feedback of the error states, either after a flight or during it.

3.2.1 Post Flight Feedback

The INS can be maintained in a calibrated state by feeding back the gyro drift estimates after a flight. These are then used to update the mini-bias trims when the INS is next used. The associated covariances from the Kalman filter can also be used

in a weighting algorithm to determine the proportion of the gyro drift estimate that is used to update the mini-biases.

3.2.2 In Flight Feedback

If error states are fed back in flight, the Kalman filter can provide an in-air alignment facility. The primary requirements of the alignment are that the platform will be level and the instrument orientation relative to true north is established. The Kalman filter estimates of tilt and heading error will be fed back to the INS which will then use them to level the platform and measure the direction of true north. The INS will also make use of GPS position and velocity data to keep the platform level during the alignment process.

The Kalman filter will have different weighting associated with the error states during the alignment compared with the normal (in-flight calibrate) mode.

3.3 USE OF SPS DATA

The behaviour of the system when using SPS data needs to be investigated further. This has already been addressed to some extent because the initial GR7 flights had to use SPS data. The measurement noise values used by the Kalman filter were increased to take account of SA and the accept/reject limits for the use of the GPS data were also increased. No attempt was made to look at the characteristics of the SA and optimise the filter for the use of this kind of data.

The effect of the SA was very noticeable, especially on the velocities. The oscillation due to SA can clearly be seen when the INS and GPS velocities are differenced. The Kalman filter currently follows the oscillations due to SA in position and velocity which may account for the slow settling of the filter that has been observed.

4 CASE STUDY 2-FIN 1010G

A second case study is the proposed retrofitting of GPS into the Tornado as a Mid-Life Improvement.

A study has been conducted jointly by GEC Ferranti and GEC Plessey to incorporate a GPS receiver card set within a modified FIN1010 INS to provide an enhanced navigation capability. There are several distinctions between the proposed approach and that adopted for the initial phase of FIN1075G, resulting from the differing requirements of the programmes.

4.1 FIN 1010G ARCHITECTURE

The FIN1010G design incorporates an inertial platform, INS electronics, a GPS receiver card set and an additional 68020 processor which hosts the

integration software. The GPS and INS Subsystems are capable of completely independent operation from separate power supplies and interfaces and do not share any common modules or single point failure mechanisms.

The embedding of the GPS, INS and integration filter within a single LRU allows low latency, secure, data transfer between the subsystems. Figure 4.1 shows the internal arrangement of the subsystems within FIN1010G and the outputs available to the Tornado aircraft avionics system.

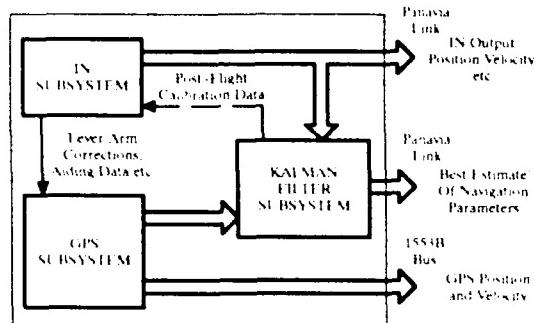


FIGURE 4.1 FIN 1010G ARCHITECTURE

4.1.1 Advantages of FIN1010G Architecture

The FIN1010G integrated GPS/INS set is specifically designed to provide the maximum possible benefit in terms of integrated performance and functionality whilst imposing minimum impact on the existing Tornado navigation sub-system architecture.

The benefits of this design solution include the following:-

- (1) Automatic INS error estimation and post-sortie updating - avoiding the need for periodic re-calibration
- (2) Significantly improved INS performance - even in flights where GPS is not available.
- (3) Enhanced GPS availability/anti-jamming performance through velocity aiding.
- (4) Use of GPS satellite range and range-rate data directly in the integration Kalman Filter allows Navigation performance to be enhanced by using data from any available satellites and does not require a full GPS solution to be available.
- (5) There is great flexibility and redundancy in adapting the aircraft avionics through the

provision of "raw" IN, "raw" GPS and the integrated INS/GPS solution.

- (6) The weight of the FIN1010G is significantly lower than the combined weight of the independent INS and GPS systems.
- (7) There is no additional volume required in the aircraft to accomodate the GPS receiver.
- (8) Aircraft installation requirements are minimised.

4.1.2 FIN1010G Integration Filter

The INS error model used in the FIN1010G Kalman Filter includes the same ten error states as those used in FIN1075G integration Filter.

Since the FIN1010G inertial platform is north-slaved a number of error mechanisms remain correlated in flight with their corresponding initial tilt errors (e.g. horizontal accelerometer biases) therefore a relatively simple error model is sufficient even for long flight profiles.

4.2 QUANTITATIVE PERFORMANCE LEVELS OF FIN 1010G

4.2.1 Simulation Techniques

4.2.1.1 Simulation Program Modules

In order to estimate the performance levels of FIN 1010G in operation in Tornado aircraft a suite of simulation program modules have been employed. These modules include the following:-

- (1) GPS satellite constellation model.
- (2) GPS range and range-rate measurement error model. (including ground, space, atmospheric and user segment errors)
- (3) FIN 1010G gimballed INS sub-system error model.
- (4) FIN 1010G error coefficient ageing model.
- (5) FIN 1010G Kalman Filter model.
- (6) FIN 1010G Monte-Carlo performance analysis executive.

4.2.1.2 Simulation Rationale

The FIN 1010G system will provide a number of performance improvements, operational advantages and additional functionality to the Navigation sub-system of the Tornado aircraft. These have been described qualitatively in Section 4.1 above. In order to quantify the anticipated performance improvements a series of detailed

simulations have been performed – the results of which are presented in Section 4.3 below.

The performance of the FIN 1010 INS currently fitted to the Tornado has also been simulated over the same operational scenario to enable direct comparisons to be made.

4.2.1.3 FIN 1010 Performance Simulations

FIN 1010 performance has been simulated over the long-duration flight profile defined in the Production Specification for the Inertial Navigation Equipment for the Tornado aircraft.

In general the performance of the FIN 1010 INS becomes gradually poorer with increasing time interval from the most recent calibration update due to long term changes in instrument error coefficients.

Recalibrations are therefore performed as required to maintain good performance – with a typical recalibration interval of around one year.

To take into account these changes with time and variations between different production INS equipments, a set of FIN 1010s have been modelled. Error coefficients of each of these INSs have been generated such that for each of the ten INSs they are known at defined intervals from the time of delivery.

The statistical properties of the error coefficients after calibration and their ageing characteristics have been determined and verified by analysis of calibration trends observable from field reports and long term inertial instrument test data. In order to take account of the “on-condition” recalibration process predicted position error is examined following each simulated flight. Recalibration corrections are applied when the terminal error exceeds a value equivalent to the limits used in the field to determine the need for recalibration of FIN 1010 equipments (three miles error after two hours).

Statistics of the performance of the set of 10 FIN 1010s have been produced for operation over notional “test” flights following the chosen flight profile at monthly intervals during the 12 months from delivery.

An extract from this data is presented below in Section 4.3 and demonstrates the degradation in expected performance as the simulated FIN 1010 population ages.

4.2.1.4 FIN 1010G Performance Simulations

FIN 1010G performance has also been simulated over the long duration flight profile defined in the

Production Specification for the Inertial Navigation Equipment for the Tornado aircraft.

In order to allow direct comparison with FIN 1010 performance it has been assumed that a set of 10 FIN 1010G equipments is available utilising a set of INS sub-systems of identical instrument error coefficient values (at all times) to those used in the FIN 1010 simulations described above. During normal operation with GPS sub-system data available FIN 1010G produces estimates of the major instrument error coefficient values which are subsequently used to update the calibration of the INS sub-system. This process is precisely modelled in the simulation software. During normal operation with GPS data available the internal FIN 1010G Kalman Filter produces best estimates of aircraft velocity and position at all times. These processes are also precisely modelled in the simulation software.

4.2.1.5 Operational Scenario

For simulation purposes it has been assumed that one week prior to each monthly “test” flight an identical sortie has been flown resulting in automatic calibration of the INS gyro drift coefficients. FIN 1010G performance statistics have been produced for the following outputs and operational conditions:-

- (1) Errors in velocity, position and heading outputs to the Main Computer via the existing Panavia data link for each monthly test flight and their rms over a 12 month period assuming that GPS data is not available during the test flights (but used one week earlier to calibrate the INS sub-system).
- (2) Errors in “best estimates” of position, velocity and Heading output on the additional Panavia Data link for use by the aircraft system.

4.3 SIMULATIONS RESULTS

4.3.1 FIN 1010 Performance

Figure 4.2(a) to 4.2(d) indicate the expected errors in Radial Position, North and East Velocities and Heading of the FIN 1010 over the operational scenario described in Section 4.2.1.3 above. These show the expected performance of the twelve month old population of FIN 1010s, the performance one month after delivery and the overall performance within the first twelve months after delivery. The degradation in performance with age is explained by long term changes in gyro drift coefficients.

Recalibrations were performed as in Table 4.1 below:-

FIN 1010 UNIT	MONTH OF RECALIBRATION
1	7
2	-
3	-
4	-
5	7, 11
6	8
7	9
8	9
9	-
10	7

Table 4.1

For reference the FIN 1010 specification limits are also shown on the Position, Velocity and Heading Error graphs.

It should be noted that the long term position and Heading errors are seen to exceed the INS Specification values. This results from the relatively infrequent recalibrations performed in practice compared with those allowed for the FIN 1010 specification.

4.3.2 FIN 1010G Performance

Two sets of results are presented here relating to FIN 1010G performance. These comprise the direct INS subsystem outputs and the Best Estimates outputs from FIN 1010G in the presence of GPS data during a flight.

In practice each FIN 1010G will operate with far more sorties per year in which GPS is available to calibrate the INS sub-system than the twelve assumed for the purposes of this simulation. Further more there will in general be a variety of different flight profiles flown. These factors can be expected to improve the automatic INS calibration process resulting in even better performance from the INS sub-system than demonstrated in these simulations.

Further 'tuning' of the internal FIN 1010G Kalman filter mechanism and noise parameters may give even more satisfactory performance than those presented below for the INS outputs by further improving automatic calibration.

4.3.2.1 FIN 1010G INS Sub-system Performance

This section includes the results of the simulations of FIN 1010G performance for the outputs to the Main Computer - without GPS available during the modelled flights.

Figures 4.3.(a) shows Radial Position Error (cep), Figure 4.3(b) shows rms North Velocity error, figure 4.3(c) shows rms East Velocity error and Figure 4.3(d) shows rms Heading error as functions of flight time - in each case the results represent overall performance within one year from delivery.

The performance is always within the current INS specification even though no routine calibrations have been performed.

This performance level will be maintained indefinitely in service without recalibration since the "settled" performance level is established within the first two months after delivery.

4.3.2.2 FIN 1010G Integrated System Performance

This section includes results for the FIN 1010G "Best Estimates" of Position, Velocity and Heading on the new Panavia data link outputs to the aircraft system - with GPS available during the modelled flights.

Figure 4.4(a) shows Radial Position Error (cep), figure 4.4(b) shows rms North Velocity error, figure 4.4(c) shows rms East Velocity error and Figure 4.4(d) shows rms Heading error as functions of flight time. In each case the results show the overall performance within one year from delivery.

The performance is always well within the INS specification even though no routine calibrations have been performed. This performance will be maintained indefinitely in service without recalibration since the "settled" performance level is established within the first two to three months after delivery.

Comparison with the results presented in Section 4.3.1 above show the dramatic improvements in performance obtainable with a fully integrated GPS/INS navigation system. This quality of performance is obtainable from FIN 1010G without any modification of the existing Tornado aircraft avionics.

5 CONCLUSIONS

5.1 RETROFIT CONSIDERATIONS

Improvements in GPS coverage and the increasing availability of user equipment offers the potential for significant enhancement of existing aircraft navigation systems. This in turn extends the operational capabilities of the aircraft. In order to provide the maximum availability and performance, the GPS system must be integrated with other navigation sensors.

There are as many possible ways of integrating GPS within a navigation suite as there are different suite

designs. The two case studies in this paper have shown some of the difficulties associated with fitting GPS to an aircraft with a modern avionics fit and have described how they may be overcome.

5.2 SINGLE CARD INTEGRATION AND GPS INTERFACE

The two case studies gave two examples of how an existing INS can be modified to interface with the GPS and make use of its data without requiring modifications to the rest of the navigation suite. This is not necessarily the optimal way to fit the GPS but it certainly is a cost effective method. It also has the advantage that the addition of the GPS is transparent to the pilot in that the operation of the INS is not altered by the presence of the GPS.

5.3 DEMONSTRATION OF SYSTEM CAPABILITY

The FIN 1075G flight trials conducted so far have shown the tremendous improvements in navigation accuracy and pilot workload reduction that can be achieved by the incorporation of GPS. They have also demonstrated a rapid reaction capability by making the navigation accuracy independent of the INS alignment time, thus giving increased operational flexibility.

Further flights will also demonstrate a reduction in the cost of ownership due to the constant calibration of the INS.

In the case of FIN 1010G, simulations have demonstrated the potential for long term navigation accuracy exceeding the performance

capability of stand alone GPS coupled with the high bandwidth of inertial navigation data.

5.4 FURTHER EXPANSION

The systems described can be expanded to provide a variety of extra facilities such as aided alignment and INS performance monitoring for maintenance purposes.

The GPS aided alignment gives the aircraft a rapid reaction capability removing any uncertainty in the initial position and allowing the alignment to proceed during taxi and take-off. It can also be used to give an autonomous at-sea alignment capability thus increasing the operational flexibility of the aircraft.

The error estimates from the Kalman filter can be used not only to calibrate the INS but also to monitor trends in the errors and give warning of an incipient INS failure before it has an impact on the mission.

6 ACKNOWLEDGEMENTS

The authors wish to acknowledge the dedicated efforts of colleagues at GEC Ferranti and GEC Plessey in conducting this work, and especially appreciate the work done by the Ministry of Defence and DRA in assessing the trials results from the FIN 1075G exercise described in Section 2.

The efforts of the pilots and the Special Projects Team of the SAOEU who conducted the flight trials of the FIN 1075G on the Harrier GR7 are also gratefully acknowledged.

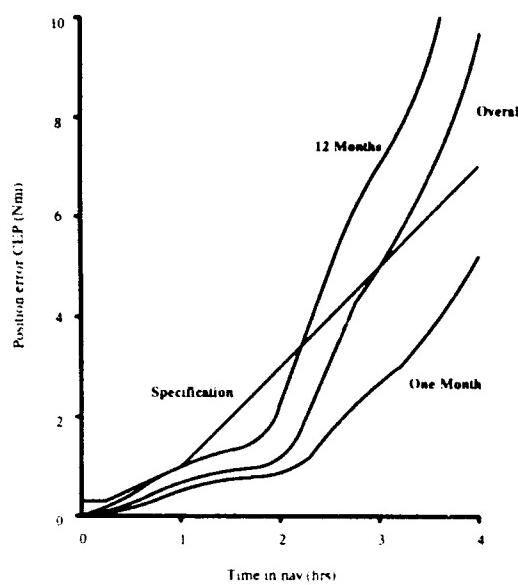


FIGURE 4.2 (a) FIN 1010 INS - RADIAL POSITION ERRORS

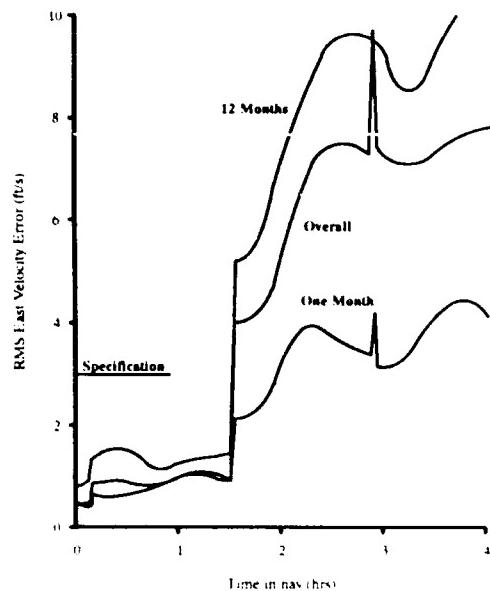


FIGURE 4.2 (c) FIN 1010 INS - EAST VELOCITY ERRORS

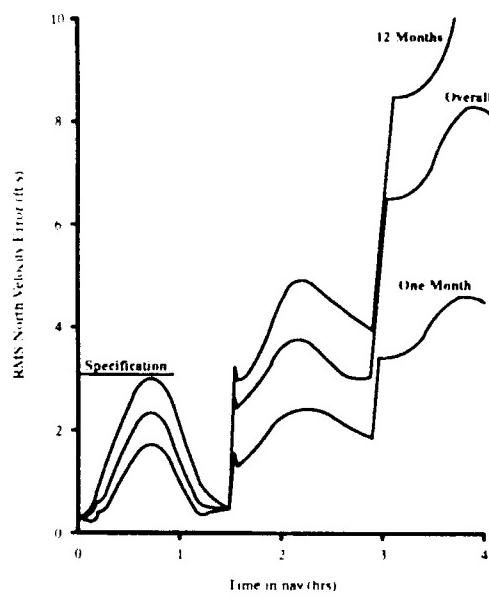


FIGURE 4.2 (b) FIN 1010 INS - NORTH VELOCITY ERRORS

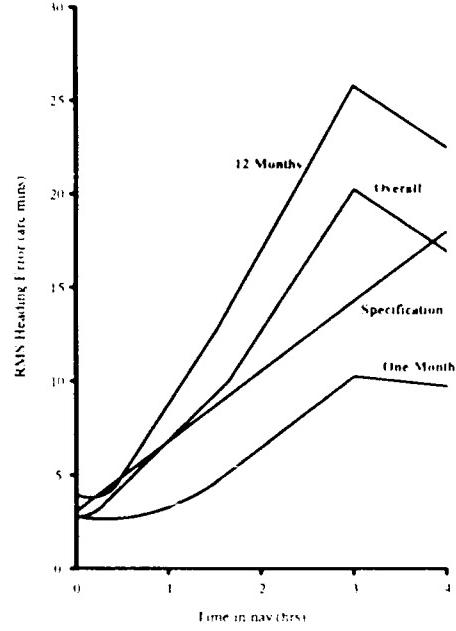


FIGURE 4.2 (d) FIN 1010 INS - HEADING ERRORS

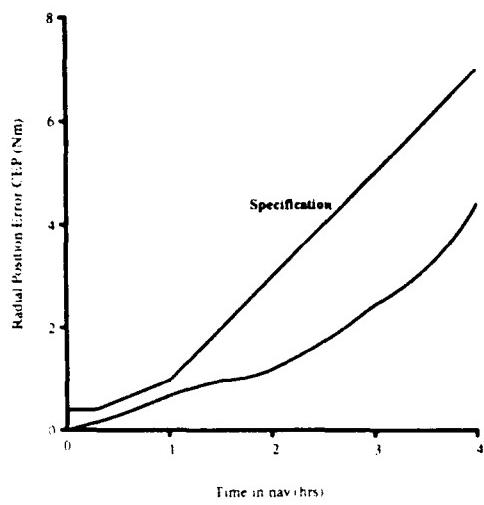


FIGURE 4.3 (a) FIN 1010G INS - RADIAL POSITION ERRORS

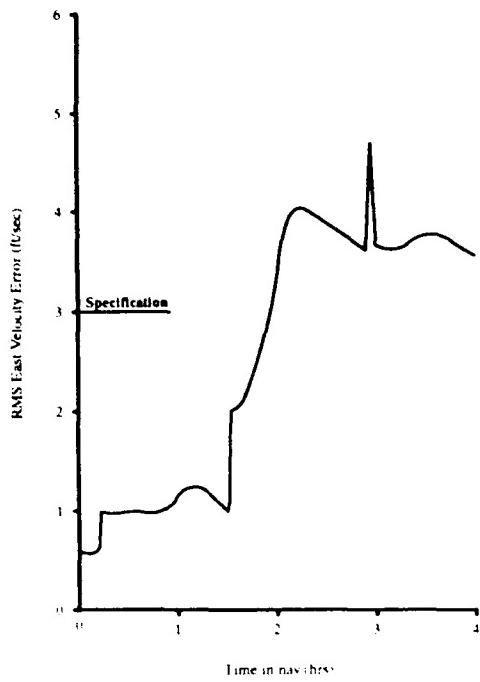


FIGURE 4.3 (c) FIN 1010G INS - EAST VELOCITY ERRORS

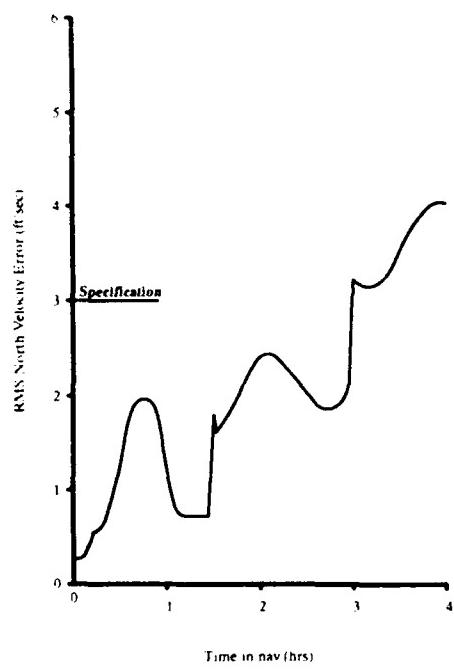


FIGURE 4.3 (b) FIN 1010G INS - NORTH VELOCITY ERRORS

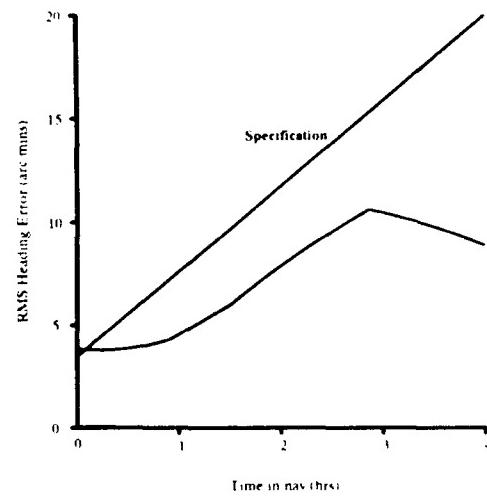


FIGURE 4.3 (d) FIN 1010G INS - HEADING ERRORS

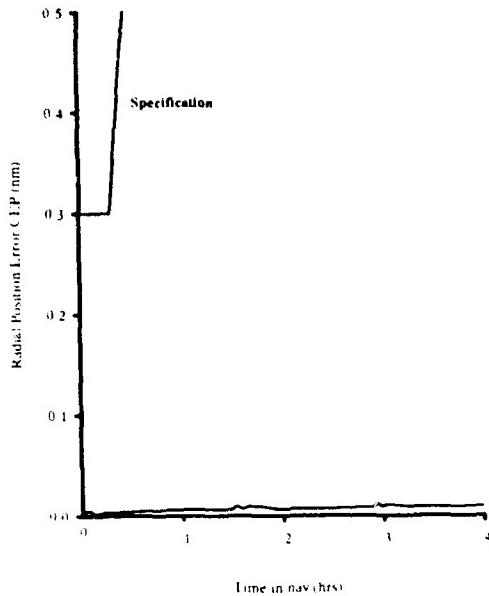


FIGURE 4.4 (a)
FIN 1010G - GPS/INS RADIAL POSITION ERRORS

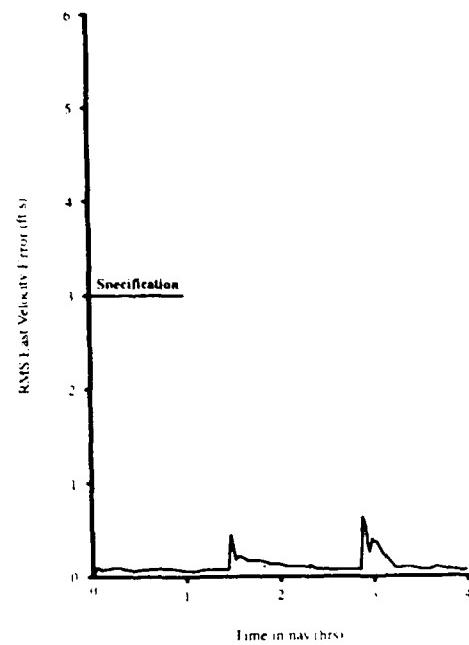


FIGURE 4.4 (c)
FIN 1010G - GPS/INS EAST VELOCITY ERRORS

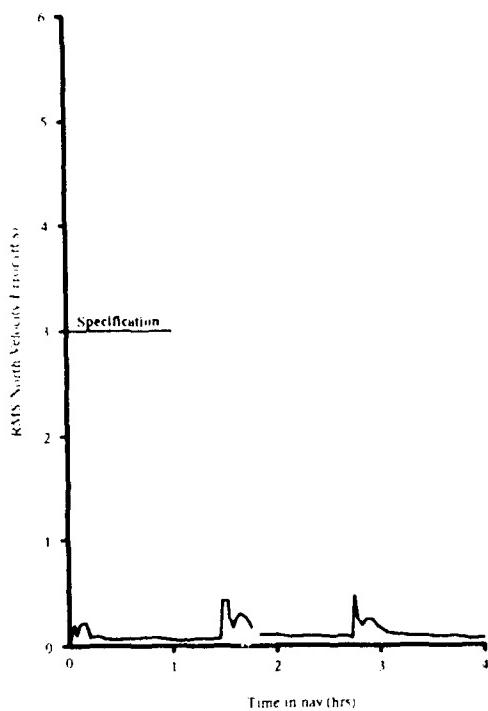


FIGURE 4.4 (b)
FIN 1010G - GPS/INS NORTH VELOCITY ERRORS

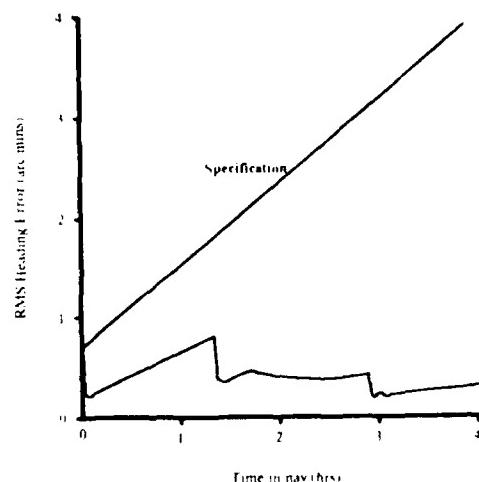


FIGURE 4.4 (d)
FIN 1010G - INS/GPS HEADING ERRORS

Integrated Precision Navigation System

**Prof. Dr.-Ing. G.Schänzer
Dipl. Ing. B.Tiemeyer**

Institute of Flight Guidance and Control
Technical University of Braunschweig
Hans-Sommer-Str. 66
D-3300 Braunschweig
Germany

Abstract

Combined Satellite and Inertial Navigation Systems can achieve extremely high positioning accuracies in the submeter range even in the dynamic environment of aircraft.

This paper presents the concept of the "Integrated Navigation System" developed at the Institute of Flight Guidance using coupled satellite and inertial sensors. Flight test results are shown, which demonstrate that this system has the potential to achieve the accuracy requirements according to ICAO CAT III for high precision approaches even under bad weather condition..

1 Introduction

When the American satellite navigation system GPS (fig.1) and its Russian equivalent GLONASS are fully operational, there will be two navigation systems available, which are more accurate than any other medium or long range navigation system.

Today the use of these two systems is restricted in time, because not all of the provided space vehicles are launched operational. This means that the geometrical configuration of the satellites is not sufficient for precise position determination at all times, but this will change in the near future.

At the Institute of Flight Guidance of the Technical University of Braunschweig an "Integrated Navigation System" is under development, which couples GPS in the differential mode with inertial sensors

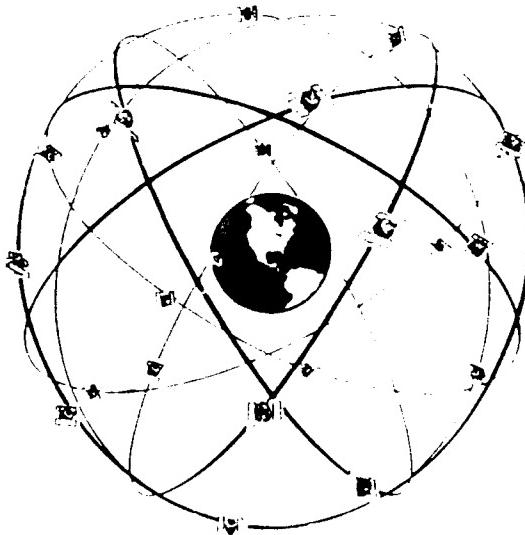


Figure 1: Global Positioning System GPS [1]

by Kalman filtering.

Using this system, it is possible to improve the position accuracy of GPS C/A-code receivers to meet the requirements for precision approach and landing even under bad weather conditions according to ICAO CAT III.

In July 1989 the first worldwide automatic landing using GPS/INS was officially demonstrated with the research commuter aircraft owned by the Institute. During several approaches to Hannover airport runway 27R the "Integrated System" was compared

to the Instrument Landing System (ILS), which is known as the European reference system in ICAO ANNEX 10 [2]. The flight test results show: The system meets ICAO CAT III accuracy requirements.

When these satellite navigation systems are available worldwide 24 hours a day, precision approaches and landings can be performed at any place in the world not equipped with Instrument Landing Systems (ILS) or Microwave Landing Systems (MLS). This will be a great advantage especially for General Aviation operating from regional airfields, which are not equipped with such expensive systems like ILS or MLS. Approaches under bad weather conditions become possible. Regularity and flight safety will increase significantly.

Several applications and tests at the Institute of Flight Guidance show: This "Integrated Navigation System" will allow the guidance of an aircraft during all phases of the flight, like taxiing, take off, enroute, approach and landing.

1.1 Navigation Systems in Use Today

Today aircraft are equipped with a variety of navigation systems depending on the application (tab 1).

Application	Range [km]	System	Required Accuracy [m]
Long Range Navigation	10,000	INS VLF-Omega	20,000
Medium Range	500	VOR/DME	200
Precision Approach	50	ILS/MLS	2

Table 1: Navigation Aids Today

In general the required accuracy of a system decreases with the required range. For long range navigation, aircraft are equipped with Inertial Navigation Systems (INS) and Very Low Frequency Omega, where the horizontal position of the aircraft is determined by a hyperbola method. Height is measured by radar, radio, or barometric altimeters.

For medium range navigation VOR (Very high fre-

quency Omnidirectional Radiorange) and DME (Distance Measurement Equipment) are used. A combination of both allows to determine the horizontal position.

Precise approach and landing are tasks which require a high accuracy, not only in horizontal position but also in height. These tasks are fulfilled by Instrument Landing Systems (ILS) and Microwave Landing Systems (MLS).

A comparison of the accuracies of the different navigation systems is shown in figure 2.

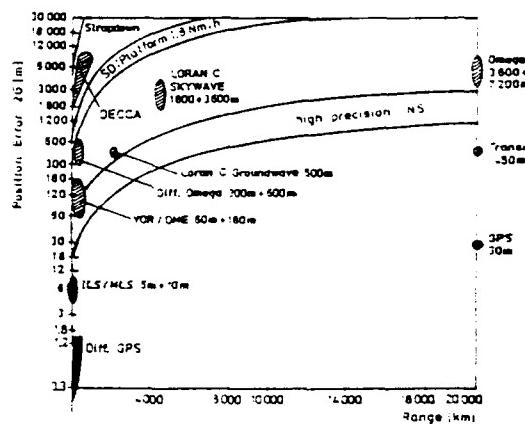


Figure 2: Position Errors Versus Distance for Different Positioning Systems [3]

2 ILS Accuracy Requirements

The accuracy requirements for the ground equipment of Instrument Landing Systems (ILS) are defined in ICAO ANNEX 10 [2]. For a horizontal visibility of 200 m and a vertical visibility of 0 m the ILS equipment has to fulfill the requirements for ICAO CAT III. For a standard 3000 m runway with a 3 degree glide path the guide beam characteristics are described in figure 3. The deviation of the aircraft due to errors of the ILS has to be less than 4.5 m horizontal and 1.2 m vertical.

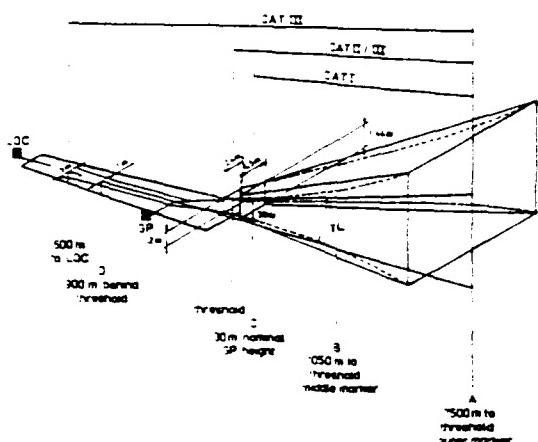


Figure 3: ILS Tolerances

3 The Use of GPS

For about two decades the American GPS has been under development. The specified accuracy for military use is in the range of 15 m (PPS: Precise Positioning Service). For civil users a reduced accuracy of 100 m will be provided (SPS: Standard Positioning Service). This accuracy will be available worldwide 24 hours a day. It presents a big step forward in civil aviation. One system can be used for long and medium distances as well as for terminal flights. The main principle of position determination with this satellite system will be described briefly (fig 4). The signals transmitted by the satellites include the parameters (ephemeris) from which the trajectories of the satellites can be calculated. Using atomic clocks onboard the satellites and precise crystal oscillators in the receiver the propagation time of the signals can be measured. Then the pseudoranges are computed by multiplying the propagation time by the speed of light. This range includes the range

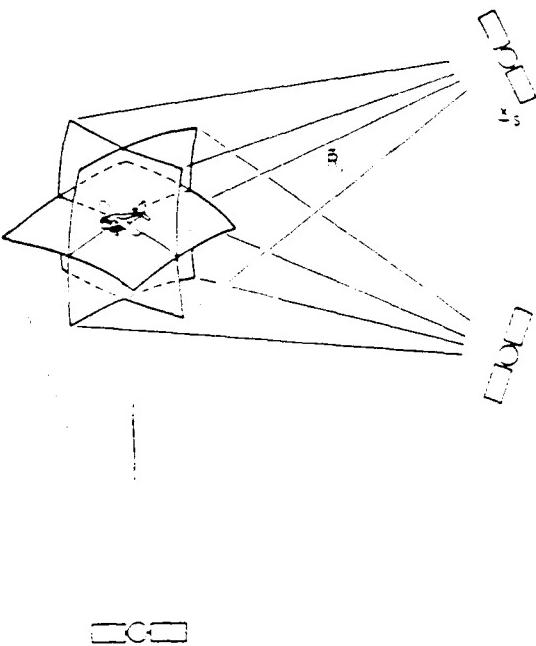


Figure 4: Basic Principle of GPS

error produced by the time offset between the receiver clock and the synchronized satellite clocks. This means that four unknowns must be determined:

- latitude φ ,
- longitude λ ,
- altitude H and
- clock offset $\Delta\tau$

For the calculation of a three-dimensional position it is necessary to use measurements from at least four satellites in sight.

3.1 GPS in Differential Mode

A typical position error of a stationary operated GPS receiver is given in figure 7. Using the raw data of a 10 channel C/A-code receiver the position calculation was done with code and phase measurements during an observation time of approximately 3 hours.

The corresponding Geometric Dilution Of Precision (GDOP) and the number of satellites (Comp-SV) which are used for the computation are depicted in figure 5.

In figure 7 the effect of the artificial reduction of the satellite signal accuracy by Selective Availability (S/A) can be observed. During the time shown

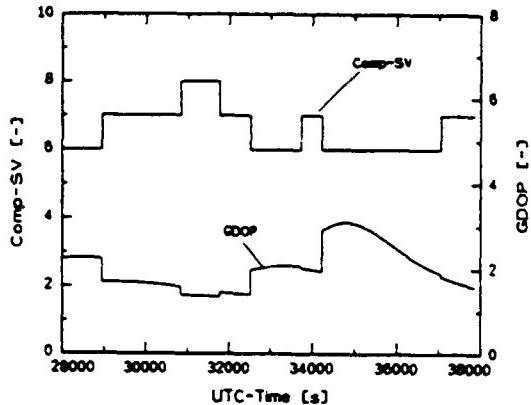


Figure 5: GDOP and number of satellites used for the position computation

in the plot the accuracy of the horizontal position (north/east) is degraded to ± 60 m in maximum. The error in the vertical position shows a larger amplitude of up to 80 m. These artificial errors superpose the natural error sources like ionospheric or tropospheric influence on the propagation of the signals or inaccuracies of the satellites' clocks.

All these errors, artificial or natural, can be eliminated by using GPS in the differential mode. In the surrounding of a known ground station we can assume that the measured errors are the same as those at an onboard receiver (Gu [4]). The transmission of the ground error data to the user (fig.6) for correction of the position calculated onboard can reduce the position error significantly.

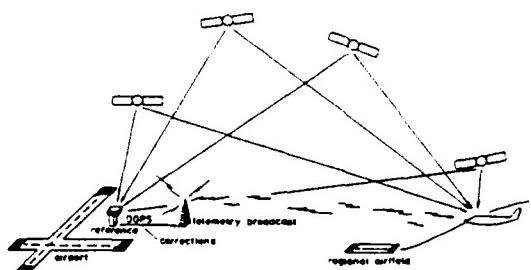


Figure 6: Differential GPS for Precision Approach Guidance

Figure 8 shows the result of the differential computation using a second receiver to correct the measurements of the first receiver (fig.7).

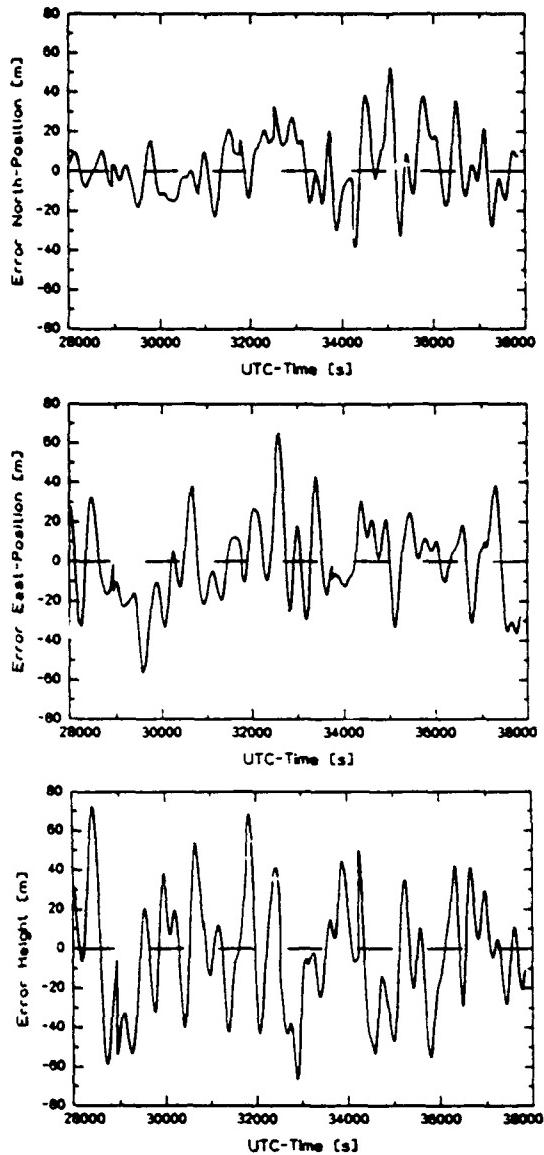


Figure 7: GPS Measurement ($v=0$)

The position errors in all three directions are reduced to the centimeter level. The observable jumps are caused by the changes of the selected satellite configuration (compare fig.6). These excellent test results are only achievable in a stationary or quasi-stationary environment. During dynamic manoeuvres tested receivers showed significant individual errors.

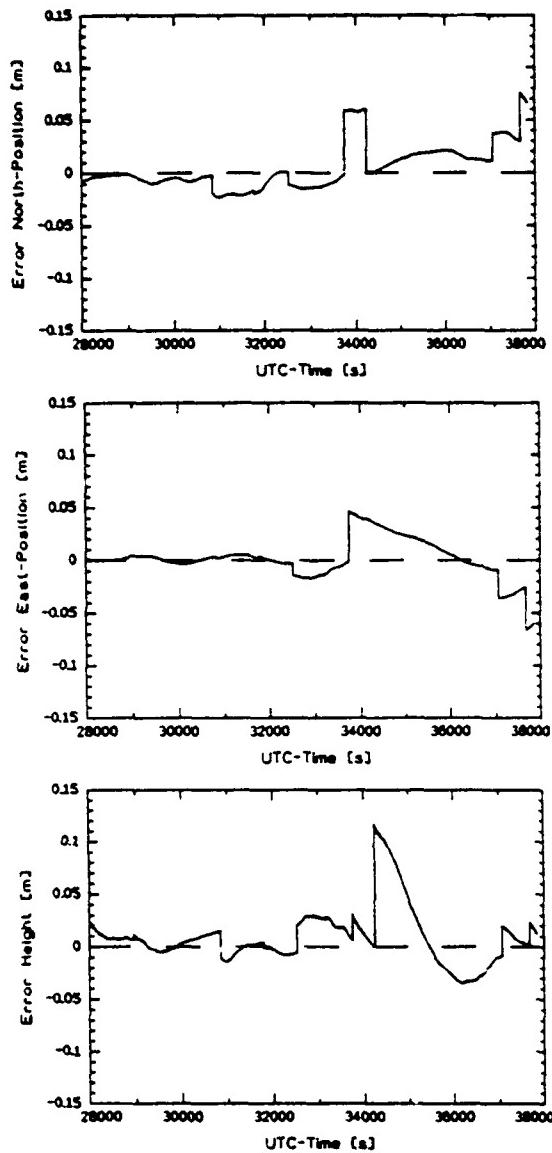


Figure 8: Differential-GPS Measurement ($v=0$)

3.2 Integrity Problems

Excellent results such as the ones mentioned above can only be achieved if the vehicle velocity is fairly constant. For a receiver moving at varying speed the bandwidth needs to be enlarged. This leads to an increase of the noise level. During manoeuvres several receivers may show time delays due to influences of acceleration on the receiver's clock and

some show a rather strong amplitude attenuation which leads to navigation errors.

For aviation, system integrity is of outmost importance. The GPS consists of four segments, the control segment, the space segment, the user segment, and the propagation path of the signals. Failures of the control segment can be avoided by parallel redundancy. They will not lead to an outage of the whole system but only to a degradation of performance caused by the use of old ephemeris and correction data. For the space segment three spare satellites are planned to ensure integrity. Outages can occur nevertheless if the visible satellites have a poor constellation. Whereas integrity of the user segment can be ensured by using several receivers, there is little to nothing one can do to ensure the integrity of the signal propagation between satellite and receiver.

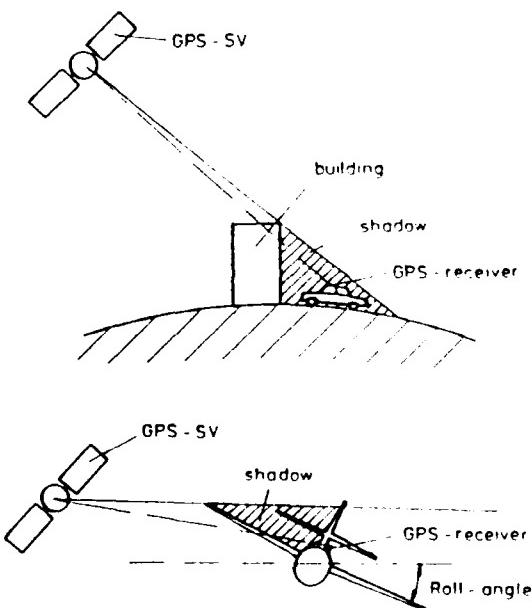


Figure 9: Masking of the GPS antenna

One type of failure of this kind is the masking of the GPS antenna during turn flights, see figure 9. In most cases, the receiver will loose contact to several and in many cases to all satellites. Total breakdowns of the position determination of up to two minutes have been observed.

When using carrier phase data unrecovered cycle slips represent outages in the sense that the desired performance cannot be achieved. For DGPS systems, the data link presents additional opportunities for failure.

For precision approaches of civil aircraft using SPS (Standard Positioning Service / C/A-Code) the obtained accuracy and integrity only by using pseudorange measurements is not sufficient. For precision applications the use of GPS in differential mode based on carrier phase measurements can reduce the errors significantly. But it becomes clear that the standard integrity requirements for air navigation and especially for precision approach cannot be fulfilled by GPS itself. The only way to improve integrity is to use GPS with a complementary sensor system like an Inertial Measurement Unit (IMU).

4 Integrated Navigation System

If the advantages of GPS are to be used without neglecting the requirements of integrity and good dynamic behaviour, a system with complementary characteristics has to be added. As seen from table 2 an Inertial Measurement Unit is an ideal candidate.

	GPS	IMU
Self Contained	no	yes
Initialization Required	no	yes
Error Behaviour	rel. high noise, stationary	low noise, drifts
Dynamics	time delay, amplitude attenuation	no noticeable time delay or attenuation
Output Rate	0.5 - 1 Hz	up to 100 Hz

Table 2: Comparison GPS - IMU

The basic idea behind the integration of GPS and IMU is to estimate the inertial sensor errors online using GPS. At the same time the IMU can be used to bridge cycle slips and times of loss of lock, but most important is to bridge the time between two

GPS position computations.

Using a Kalman filter the fully Integrated Navigation System is shown in figure 10.

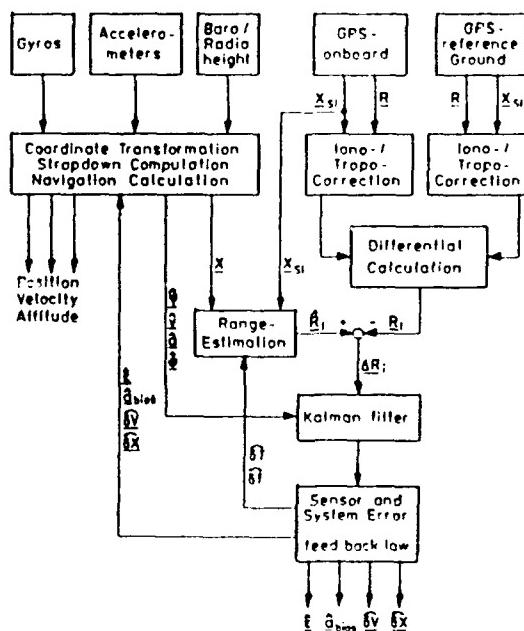


Figure 10: Fully Integrated Navigation System

On the basis of theoretical studies, simulations and flight tests took place at the Institute of Flight Guidance and Control of the Technical University of Braunschweig. For these tasks it is possible to use the two flight simulators operated by the institute, one primarily for basic research and the other one for routine tasks. In the basic research simulator a fully dynamic model of a wide body aircraft is implemented. Models for the orbital trajectories of the satellites and for the propagation of the electro-magnetic waves through the atmosphere are available. Investigations with the pilot in the loop are possible as well as the variation of typical sensor behaviours.

These tools are used in a closed chain with the test aircraft. Simulation results are validated by flight tests and the flight test results then can be understood by theory and simulator runs. For this purpose the Institute of Flight Guidance owns two research aircraft of the types Dornier DO 28, DO 128. These aircraft are fully equipped with air-data sensors, INS, GPS, telemetry, interface computers, sufficient data storage capacity, actuators and displays (fig.11).

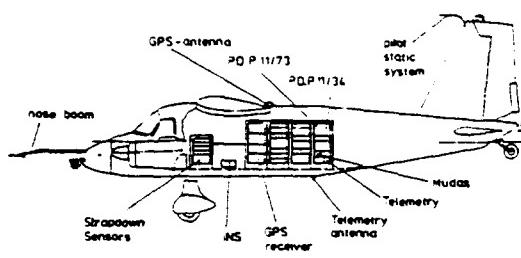


Figure 11: Cross Section DO 28

5 Applications and Examples

5.1 High Precision Approach Guidance

An integrated GPS-IMU system for precision approach and landing has been developed at the Institute of Flight Guidance at the Technical University of Braunschweig. A block diagram of this system is shown in figure 12.

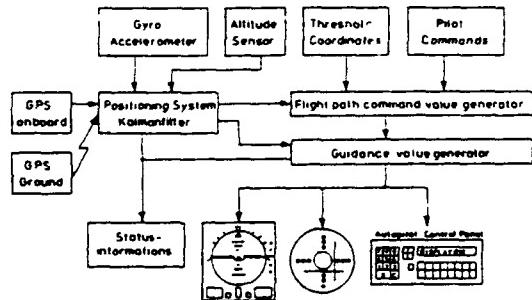


Figure 12: Integrated Flight Guidance System

The position and attitude of the aircraft is determined by Differential GPS (DGPS), IMU, and an altimeter. From the threshold coordinates, the pilot's commands, and the present position a commanded trajectory is computed and then compared to the actual trajectory flown. The difference is either displayed on an Horizontal Situation Indicator (HSI) and the cross pointer, or input to the autopilot which computes commands for moving the control surfaces.

This system allows even steep and curved approaches. The only ground equipment needed is a GPS monitor station with a data link to transmit the differential corrections to approaching aircraft.

Using such a system the first worldwide fully auto-

matical landing was performed in July 1989 with the twin engined test aircraft Dornier Do 28 of the Technical University of Braunschweig. This took place during the public demonstration flights on July 11, 1989 at the symposium 'Satellitennavigation in der Flugführung' of the German Institute of Navigation (DGON) in Braunschweig.

Figure 13 shows the result of a flight test approaching Hanover airport runway 27R. The Hanover ILS is named as a reference system for Europe in ICAO Annex 10 [2].

The position information of the integrated navigation system is transformed to the polar coordinate system of the ILS and compared with the deviations given by the cross pointer of the ILS. Depicted is the difference between both systems versus the distance to the threshold. The borders of the boxes represent the accuracy requirements of a precision approach according to ICAO Category III. This means that the pilot has no vertical or horizontal sight until touch-down. From this point on the horizontal sight has to increase to a minimum of 200m.

The figure shows that these accuracy requirements are achieved in all cases. Only in the immediate surroundings of the threshold the differences are larger due to errors of the ILS, caused by passing the glideslope transmitter. Especially for the localizer, the results are well within the required limits. Therefore one could also imagine to use an integrated DGPS-IMU for the calibration of ILS in the future. At the moment, this is done by using a theodolite equipped with a laser tracker.

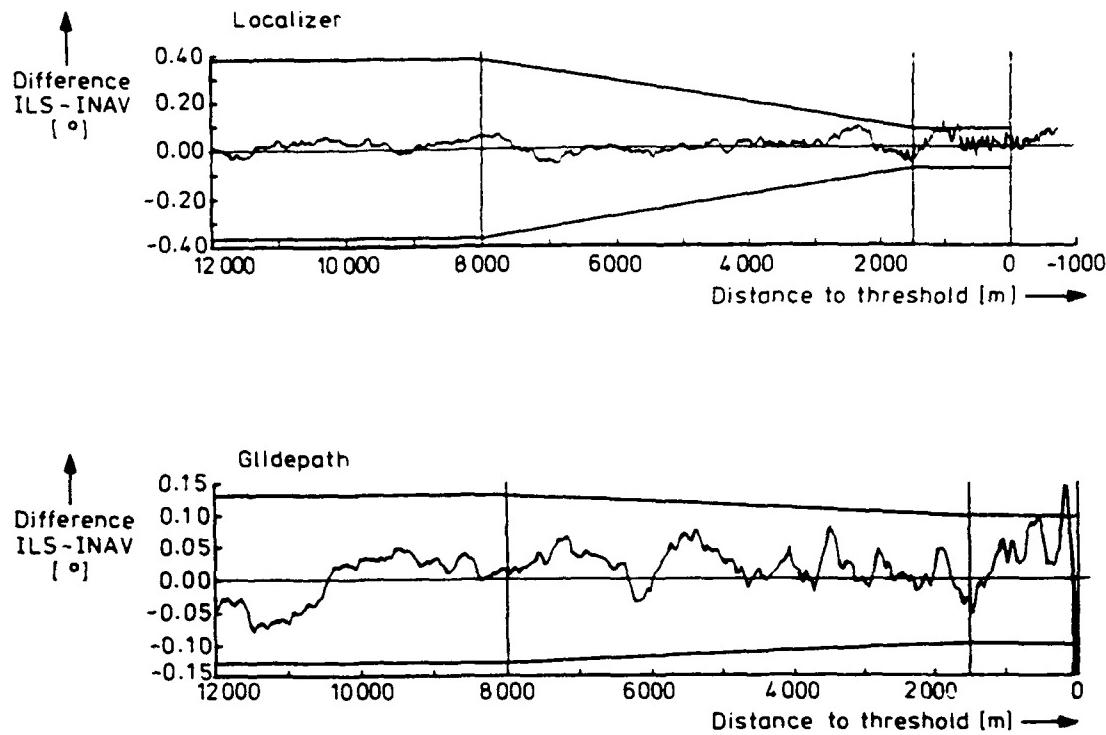


Figure 13: Inspection Flight Results for Hannover ILS 27R

5.2 Taxi Guidance

Today safe landings can be performed even under bad weather conditions using ILS. There are only few and inadequate installations though that assist the pilot to find the right way to the ramp, especially in darkness, rain, or fog. The safety and capacity of airfields could be improved if the position of the aircraft on the field would be known exactly, monitored by the pilot and controller, and if taxi instructions would be transmitted and executed automatically.

The Institute of Flight Guidance of the Technical University of Braunschweig has developed a prototype Taxi Guidance System (Möhlenkamp [5]). It is based on DGPS-IMU. The position information can be shown on a display. The airport coordinates are stored on a transputer system which is also used to identify the runway number and for the calculations of the free programmable graphics. The system has been demonstrated during the field session of the First International Symposium on DGPS in Braunschweig in September 1991.

5.3 Special Projects

Special projects where the aircraft is a platform for survey equipment, impose high requirements on positioning, attitude determination, and guidance. Such applications are e.g. aerial photogrammetry, remote sensing, or airborne gravimetry.

Aerial photogrammetry requires that the pictures taken overlap to a certain percentage. To achieve this the aircraft has to be guided on a special pattern (fig. 14) with an accuracy of a few meters. Especially the intercept on the return flight is a very demanding task. With the DGPS-IMU system and a special flight director parallel stripes with a distance of 50 m can be guided with an accuracy of 5 m at moderate turbulence.

Laser profiling for contour maps requires knowledge on position and attitude of the aircraft, see figure 15.

In a test conducted in cooperation with the Institute of Photogrammetry of the Stuttgart University, heights obtained by DGPS-IMU and a laser altimeter were compared with photogrammetrically determined heights. For approximately 17.000 measurements the differences were less than 25 cm (tab.3).

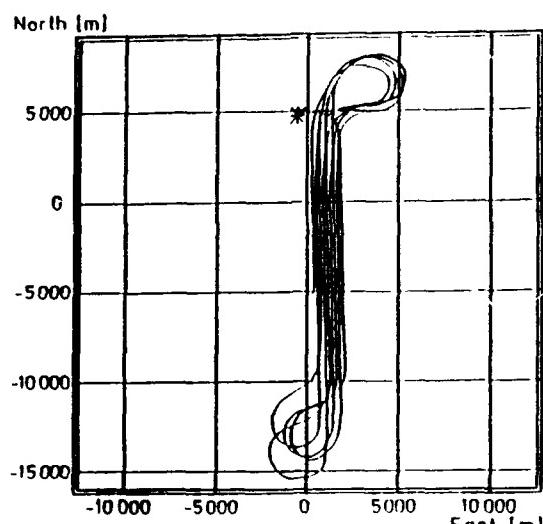


Figure 14: Ground Track of the Integrated Navigation System for Aerial Photogrammetry

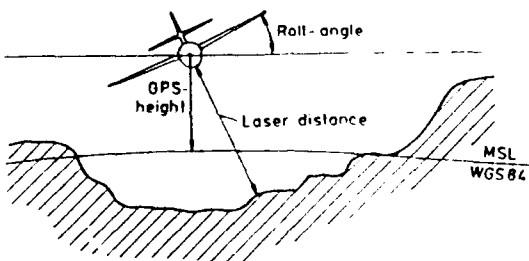


Figure 15: Earth Surface Contour Measurement

6 Summary

The Global Positioning System provides a positioning accuracy that has never been available so far from one system. GPS could replace the variety of systems used so far for enroute navigation and at the same it could replace the expensive systems for high precision approach guidance like ILS and MLS. GPS has revolutionized positioning and surveying on land. Civil aviation that is not subject to regulations and certification has already made extensive use of GPS. Thanks to GPS many tasks can now be performed at a fraction of the cost and time of conventional procedures. Other applications have only been made possible by GPS.

Provided that the availability of GPS and also of the

Height Differences between Laser Profiling and Photogrammetric Height Determination	Number of Differences	Mean	Standard Deviation
1) all controlled points	16916	-0.01 m	0.27 m
2) extreme differences (> 0.77 m) excluded	16716	-0.02 m	0.24 m
3) points in fields	15055	-0.03 m	0.24 m
4) points in forest	1661	0.05 m	0.27 m
5) western part, slope up to 15 %	4769	-0.09 m	0.26 m
6) middle part, slope up to 7 %	11353	0.00 m	0.23 m
7) eastern part, steep terrain, average slope 28 %	1443	0.03 m	0.25 m

Table 3: Comparison of Laser Profiling to Photogrammetric Control [6]

russian GLONASS will be ensured, a similar evolution will take place in regulated aviation. Applications which have not been possible so far such as automatic landings on grass strips or landings in accordance with ICAO CAT III on non-equipped airfields have already been demonstrated. This could be of interest especially for regional airfields, for airports in less developed countries and for special military applications. A number of other applications such as taxi guidance are being investigated. The potential of satellite navigation is immense. Increasing air traffic and increasing costs for air traffic control will not allow to leave this potential unused.

7 References

1. The NAVSTAR GPS System, AGARD-LS-161, September 1988
2. ICAO ANNEX 10 'Aeronautical Telecommunications', Fourth Edition April 1985, International Civil Aviation Organization
3. GPS Applications, Logsdon Th.S., Rockwell International Satellite Division
4. Gu, X.: Error Correction for Differential GPS with long separated Ground Station and User for Aircraft Landing, Proceedings GPS'90, Ottawa/Canada
5. Möhlenkamp, K.: Precision Navigation for Aircraft Taxi Guidance, Paper pres. at First Int. Symp. Real Time Differential Applications of the Global Positioning System, Braunschweig, September 16-20, 1991.
6. Ackermann, F.; Lindenberger, J.; Schade, H.: Kinematische Positionsbestimmung mit GPS für die Laser-Profilmessung, Reports of the DFG Sonderforschungsbereich 228 "Hochgenaue Navigation - Integration geodätischer

und navigatorischer Methoden", Stuttgart
University

7. Jacob, Th.: Integrated System Using Differential GPS and Inertial Measurement Unit, in: Schwarz, K.P. and Lachapelle, G. (eds.), Kinematic Systems in Geodesy, Surveying, & Remote Sensing, Springer Verlag, New York, 1991

FLIGHT EVALUATION OF DIFFERENTIAL GPS AIDED INERTIAL NAVIGATION SYSTEMS

B. David McNally, Russell A. Palelli, Ralph E. Bach, Jr., David N. Warner, Jr.
 NASA Ames Research Center
 Moffett Field, California, USA 94035-1000

Abstract

Algorithms are described for integration of Differential Global Positioning System (DGPS) data with Inertial Navigation System (INS) data to provide an integrated DGPS/INS navigation system. The objective is to establish the benefits that can be achieved through various levels of integration of DGPS with INS for precision navigation. An eight-state Kalman filter integration was implemented in real-time on a twin turbo-prop transport aircraft to evaluate system performance during terminal approach and landing operations. A fully integrated DGPS/INS system is also presented which models accelerometer and rate-gyro measurement errors plus position, velocity, and attitude errors. The fully integrated system was implemented off-line using range-domain (seventeen-state) and position-domain (fifteen-state) Kalman filters. Both filter integration approaches were evaluated using data collected during the flight test. Flight-test data consisted of measurements from a 5 channel Precision Code GPS receiver, a strap-down Inertial Navigation Unit (INU), and GPS satellite differential range corrections from a ground reference station. The aircraft was laser tracked to determine its true position. Results indicate that there is no significant improvement in positioning accuracy with the higher levels of DGPS/INS integration. All three systems provided high-frequency (e.g., 20 Hz) estimates of position and velocity. The fully integrated system provided estimates of inertial sensor errors which may be used to improve INS navigation accuracy should GPS become unavailable, and improved estimates of acceleration, attitude, and body rates which can be used for guidance and control. Precision Code DGPS/INS positioning accuracy (root-mean-square) was 1.0 m cross-track and 3.0 m vertical. (This AGARDograph was sponsored by the Guidance and Control Panel.)

List of Symbols

e_p	aircraft position vector in the Earth centered-Earth-fixed (ECEF) frame
e_v	aircraft velocity vector in ECEF frame
e_g	Gravity vector in ECEF frame
λ, μ, h	geodetic position (latitude, longitude, altitude)
$e_{\dot{v}}$	aircraft velocity vector in corrected level frame
e_g	gravity vector in corrected level frame
b_a	vector of specific-force measurements in the body frame
n_a^{INU}	vector of specific-forces in the INU level frame
$b_b A$	vector of accelerometer biases
K_A	matrix of accelerometer scale factors
b^{ω_T}	vector of angular rate-gyro measurements in the body frame
n_{ω_E}	vector of Earth rates in INU and corrected level frame
$n_{\omega^{INU}}$	vector of craft rates in the INU level frame
c_{ω_V}	vector of craft rates in the corrected level frame
$n_{\omega^{INU}}$	angular velocity vector of the INU level frame
c_{ω_L}	angular velocity vector of the corrected level frame
$b_b G$	vector of body frame angular rate gyro biases
K_G	matrix of angular rate gyro scale factors
c^{T_N}	matrix transformation from INU level frame to corrected level frame
n^{T_B}	matrix transformation from body frame to INU level frame
c^{T_B}	matrix transformation from body frame to corrected level frame
e^{T_N}	matrix transformation from INU level frame to ECEF frame
b_{baro}	barometric altimeter measurement
b_B	barometric altimeter bias

R_m	Earth meridional radius of curvature
R_p	Earth polar radius of curvature
Ω_E	Earth rotation rate
p	pseudorange measurement
Δr	range measurement
Δt_{int}	deltarange integration time interval measurement
e_r	receiver's estimate of total range error (less ionospheric contribution)
b_r, d_r	receiver's estimate of receiver clock bias and clock drift
Δt_{drift}	pseudorange valid time minus clock bias valid time
r, \dot{r}	estimate of range and range-rate
r_0, \dot{r}_0	computed range and range-rate from reference station to satellite
$\Delta r, \Delta \dot{r}$	differential range and range-rate corrections
Δt_{lat}	uplink latency (airborne pseudorange valid time minus differential correction valid time)
R, \dot{R}	differentially corrected range and range-rate
b, d	receiver clock bias and clock drift
$\mathbf{p}_s, \mathbf{v}_s$	vectors of satellite ECEF position and velocity derived from ephemeris

Introduction

The highly accurate world wide positioning capability of the Global Positioning System (GPS) has revolutionized navigation. Over the past ten years many studies have evaluated the use of GPS for a variety of air, sea, and land positioning and navigation applications. The precision military GPS code (P Code) provides 10-20 m positioning accuracy, the coarse acquisition code (C/A Code), the civilian GPS code, provides 20-40 m positioning accuracy (100 m when Selective Availability is activated) [1-8]. Local corrections to satellite range measurements are the basis of Differential GPS (DGPS), which provides an even more accurate solution. Positioning accuracies of 1.2 m in the horizontal plane and 2.4 m vertical have been demonstrated with DGPS [5-8,12-14].

The demonstrated accuracy of GPS has generated a great deal of interest in evaluating the use of GPS and DGPS for navigation to support terminal area flight operations, including approach and landing [2,5-14]. The integration of DGPS with an Inertial Navigation System (INS) is of particular interest. The INS provides excellent transient response but

lacks long term stability. The GPS has excellent long term stability but has limited transient response characteristics. When properly integrated, the two systems complement each other very well to realize a high frequency, high accuracy navigation system.

DGPS research at NASA Ames began in the early 1980's when a C/A Code DGPS system was developed and flight tested on an SH-3G helicopter using single channel sequencing receivers [11-14]. The objective was to evaluate the use of DGPS to support helicopter terminal approach operations. Positioning accuracies during final approach were very encouraging (5.2 ± 4.0 m (1σ) lateral and 7.7 ± 3.5 m (1σ) vertical). Vertical axis results were even better when the solution was aided with barometric (5 ± 3 m (1σ)) and radio (5 ± 2 m (1σ)) altimeters [12,14]. The time lags due to satellite sequencing, however, were prohibitively large for real-time navigation and guidance without inertial aiding. Nevertheless, the positioning accuracies attained during these early tests demonstrated the potential of DGPS for terminal approach and landing operations. Static tests indicated that 2-3 m accuracy might be achieved in flight using future generation multi-channel equipment [13].

In 1989 NASA Ames entered into a joint program with the Department of Defense and the Federal Aviation Administration to evaluate, through flight test, DGPS positioning accuracy and to determine operational procedures for making effective use of DGPS for terminal area flight operations. The program was divided into two phases. Phase 1 has evaluated the use of the Precise Positioning Service (PPS, i.e., P Code) of the GPS for approach and landing operations. Phase 2 will evaluate use of the Standard Positioning Service (SPS, i.e., C/A Code) aided with carrier phase measurements for terminal area operations with particular emphasis on the use of kinematic carrier tracking for precision navigation.

Initial results of the recent flight test evaluation of P Code DGPS INS for terminal area operations are described in [5]. A real-time DGPS INS system was installed in a twin turbo-prop transport aircraft. The navigation algorithm was a simple eight state Kalman filter integration of DGPS with INS. Positioning accuracy during approach and landing was about 1 m (1σ) horizontally and 3 m (1σ) vertically. This paper reviews the navigation

algorithm that was evaluated in [5] and then extends that algorithm to a more complete integration of DGPS with INS. The objective is to establish the benefits that are achievable through a complete integration of DGPS with INS for precision navigation. Although there are many papers in the navigation literature that have explored methods for integration of GPS with INS [6,7,9,10], validation with flight data has been limited.

The paper begins with a review of the real-time (eight state) DGPS/INS algorithm. The next section describes the eighteen-state and sixteen-state fully integrated navigation algorithms. The test equipment and the flight experiment are described next. In the results and discussion section, the fully integrated navigation algorithms are validated using the flight data base and compared to the results from the eight-state system. The paper closes with some concluding remarks. Three appendices present more detailed discussions of GPS pseudorange and deltarange measurement processing, the derivation of the tilt matrix kinematics, and the Kalman filter implementations for the fully integrated systems.

Real-Time DGPS/INS System

The block diagram shown in Fig. 1 illustrates the integrated navigation system that was implemented in real-time during flight-test evaluation of P-Code DGPS/INS for approach and landing [5].

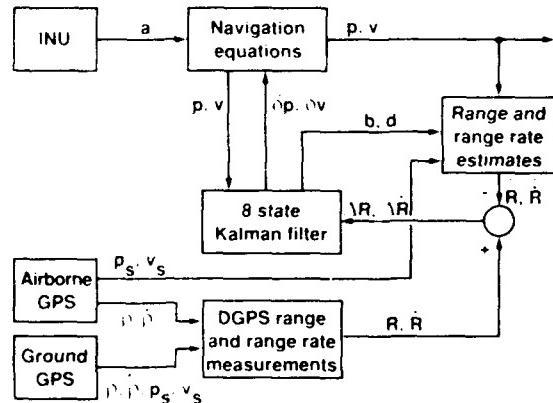


Fig. 1 - Eight-State Range-Domain Algorithm

The navigation equations were integrated at 1 Hz in an Earth-centered-Earth-fixed (ECEF) coordinate frame. An ECEF frame is convenient since satellite

position and velocity, computed from GPS satellite ephemeris parameters, are given in ECEF coordinates. This ECEF frame has the X axis pointing through the equator at the Greenwich meridian, the Y axis 90° to the East, and the Z axis pointing through the North pole. The navigation equations in the ECEF frame are given by

$$\dot{e}p = eV \quad (1)$$

$$\dot{e}V = eT^h a^{inv} + eg \quad (2)$$

Measurements from the Inertial Navigation Unit's (INU) locally-level coordinate frame were sampled at 20 Hz and stored in a buffer. A navigation state update was triggered upon arrival of a new set of measurements from the GPS receiver. Specific forces from the locally-level frame (a^{inv}) were converted to the ECEF frame and then averaged over the 1 second GPS update interval. The average specific-force was corrected for gravity and integrated twice to project position and velocity to the next (current) GPS measurement time. A flat, non-rotating Earth model was assumed so specific-forces were not corrected for Coriolis and centripetal forces. Gravity was assumed constant and set to the gravity at the test site, since approach and landing speeds were relatively slow (120 kt) and flight tests were conducted at a centralized area at low altitude. Measurements from the INU were used directly without correcting for accelerometer or tilt errors.

Pseudorange measurements from the airborne GPS receiver were first adjusted for satellite clock bias, tropospheric delay, channel calibration and hardware delays, and the receiver's estimate of clock bias. Differential range corrections, uplinked from the ground reference station, were then applied to provide the best estimate of range to the satellite. Range-rate was approximated by dividing the deltarange measurement by the time interval over which the measurement was made. Range-rate data were then adjusted with range-rate corrections from the ground station. Differential corrections were uplinked from the ground reference station and applied at 1 Hz. Processing of pseudorange and deltarange measurements and the differential correction algorithm are described in more detail in Appendix A.

Differentially corrected range and range-rate data were input (as measurements) to an eight-state Kalman filter to update position and velocity. The measurement model was based on the familiar range and range-rate equations given by

$$\rho = [(\rho_p - \rho_s)^T (\rho_p - \rho_s)]^{1/2} + b \quad (3)$$

$$\dot{\rho} = [(\rho_p - \rho_s)^T (\dot{\rho}_p - \dot{\rho}_s)]/\rho + d \quad (4)$$

The clock correction equation relating GPS receiver clock bias and clock drift (bias rate) is given by

$$\dot{b} = d \quad (5)$$

Range and range-rate estimates were computed using (3) and (4) with position and velocity state projections from (1) and (2). Note that there is no coupling between the navigation and the clock correction states. Differentially corrected range and range-rate measurements were differenced with range and range-rate estimates to drive the Kalman filter. The filter computed corrections to position, velocity, clock bias, and clock drift and applied these corrections to the navigation state vector. The 1 Hz navigation solution was then written to disk for post flight analysis.

The real-time computer generated simulated Instrument Landing System (ILS) localizer and glideslope signals using position and velocity from the real-time (1 Hz) DGPS/INS solution combined with unaided INU velocity at 20 Hz. The 1 Hz DGPS/INS position and velocity data were converted from ECEF to the local runway coordinate system (RCS), aligned with the runway and centered at the touchdown point. DGPS/INS position in the RCS frame was propagated at 20 Hz using INU velocity (converted to the RCS frame) calibrated with DGPS/INS velocity. Localizer and glideslope commands were computed as the angular deviation of the aircraft position from a nominal 3° glide path. Glideslope commands were referenced to the origin of the RCS; localizer commands were referenced to a point 7,000 ft (2,134 m) beyond the RCS origin on the runway centerline. The computer-generated guidance commands were interfaced to standard ILS-type localizer and glideslope guidance displays in the cockpit.

The navigation algorithm used data from a GPS receiver and the INU, but ran in parallel on a separate processor. This processing architecture

was required because the GPS receiver did not accept differential corrections. (The terms "INS" and "INU" are often used synonymously in the navigation literature. In this paper "INS" is the more general term, and "INU" refers to the particular unit that was used in the flight-test evaluation.)

Fully Integrated DGPS/INS Algorithms

The purpose of a fully integrated DGPS/INS navigation system is to provide an in-flight calibration (or identification) of the INU's state and instrument errors and thereby provide a precise inertial navigation solution. The block diagrams shown in Figs. 2 and 3 illustrate two complete integrations of DGPS with INS. In these integrations, body frame accelerometer and rate-gyro errors, a barometric altimeter bias error, and the INU body-to-level attitude (or tilt) error are modeled, in addition to the position, velocity, and clock errors that were modeled in the eight-state system (Fig. 1). The system of Fig. 2 utilizes a range-domain filter, similar to that used in the eight-state filter, and is driven by satellite range and range-rate measurements. The system of Fig. 3 utilizes a position-domain filter and is driven by DGPS position and velocity measurements. Both the range- and position-domain systems use the same navigation equations.

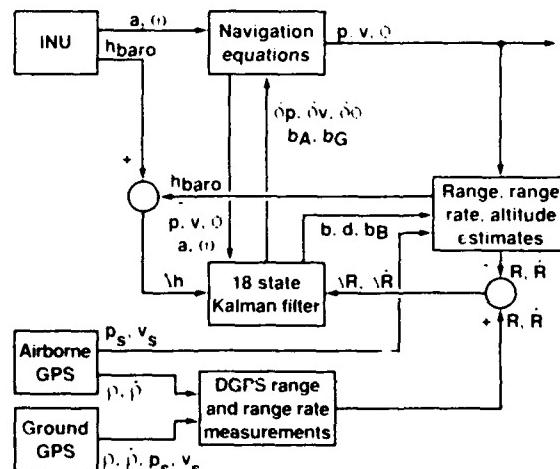


Fig. 2 - Eighteen-State Range-Domain Algorithm

Once the inertial system's instrument errors are calibrated, both of these navigation systems could

provide an accurate solution even if DGPS data become unavailable. DGPS data are unavailable, for example, between the 1 Hz samples or during receiver dropouts caused by antenna shading, jamming, or other reasons. The navigation equations were integrated at 20 Hz since the INU data were available at this rate, and 20 Hz should be adequate for terminal area operations. In an operational implementation, it is expected that the navigation computations would be done internal to the INU at a standard rate of 64 Hz or 256 Hz. The Kalman filter updates should be applied at a rate that is consistent with the dynamics of the inertial state and instrument errors and the allowable drift between updates. In this evaluation Kalman filter updates were applied at 1 Hz.

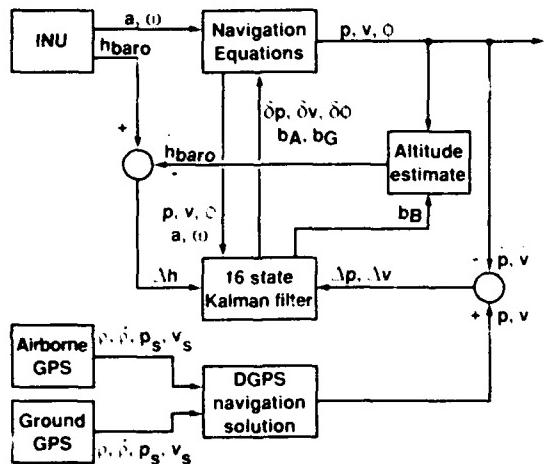


Fig. 3 - Sixteen-State Position-Domain Algorithm

Here we choose to integrate the navigation equations in a North, East, down locally-level coordinate frame. A locally-level frame is typically used for inertial navigation applications and all of the INU flight data are in a locally-level frame. The locally-level frame moves with the aircraft and the rate-gyro measurements are integrated to level the frame such that the X-Y plane remains perpendicular to the gravity vector. A North-pointing mechanization was chosen for the post-test analysis. For global usage in a real-time system, however, the wander-azimuth mechanization is more appropriate.

The navigation computations are summarized in the block diagram shown in Fig. 4. Specific-force and angular rate measurements (${}_n\alpha^{inu}$, ${}_n\omega_L^{inu}$), in the

INU locally-level frame, are the primary inputs to the navigation equations. Accelerometer and rate-gyro errors are each modeled with a bias and scale factor, although the off-line implementation uses only the bias. The need for the tilt matrix ${}_cT^n$ arises from the angular rate-gyro error which causes a "tilt" error in the INU level frame. Locally-level specific-force components from the INU are corrected for body frame accelerometer error and rotated through the tilt matrix to the corrected level frame. Specific-force components are then adjusted for gravity and Coriolis forces and integrated to give velocity in the locally-level, North-pointing frame. The vertical velocity is the rate of change of altitude. North velocity is divided by Earth's meridional radius of curvature to form latitude rate. East velocity is divided by the product of Earth's prime radius of curvature and the cosine of latitude to form longitude rate. Integration of these quantities gives current position in terms of latitude, longitude, and altitude. In this system we use the WGS-84 gravity model which computes gravity as a function of latitude, longitude, and altitude. The navigation states are geodetic position (latitude, longitude, altitude) and velocity (North, East, down), and the three components of platform tilt error. The navigation equations are

$$\dot{\lambda} = {}_cV^N/R_m; \dot{\mu} = {}_cV^E/(R_p \cos \lambda); \dot{h} = -{}_cV^D \quad (6)$$

$$\dot{{}_cV} = {}_cT^n [{}_n\alpha^{inu} - {}_nT^b (K_A b_a + b_{bA})] + {}_cG$$

$$- (2 {}_c\omega_E + {}_c\omega_V) \times {}_cV \quad (7)$$

where

$${}_c\omega_V = \begin{bmatrix} {}_cV^E / R_p \\ -{}_cV^N / R_m \\ -{}_cV^E \tan \lambda / R_p \end{bmatrix};$$

$${}_c\omega_E = \begin{bmatrix} \cos \lambda \\ 0 \\ -\sin \lambda \end{bmatrix} \Omega_E. \quad (8)$$

The kinematics required for calculating the tilt matrix are derived in Appendix B and the result is presented here as

$${}_cT^n = - \{ {}_c\omega_L \cdot {}_cT^n [{}_n\omega_L^{inu} - {}_nT^b (K_G b_{\omega_T} + b_{bG})] \} \times {}_cT^n \quad (9)$$

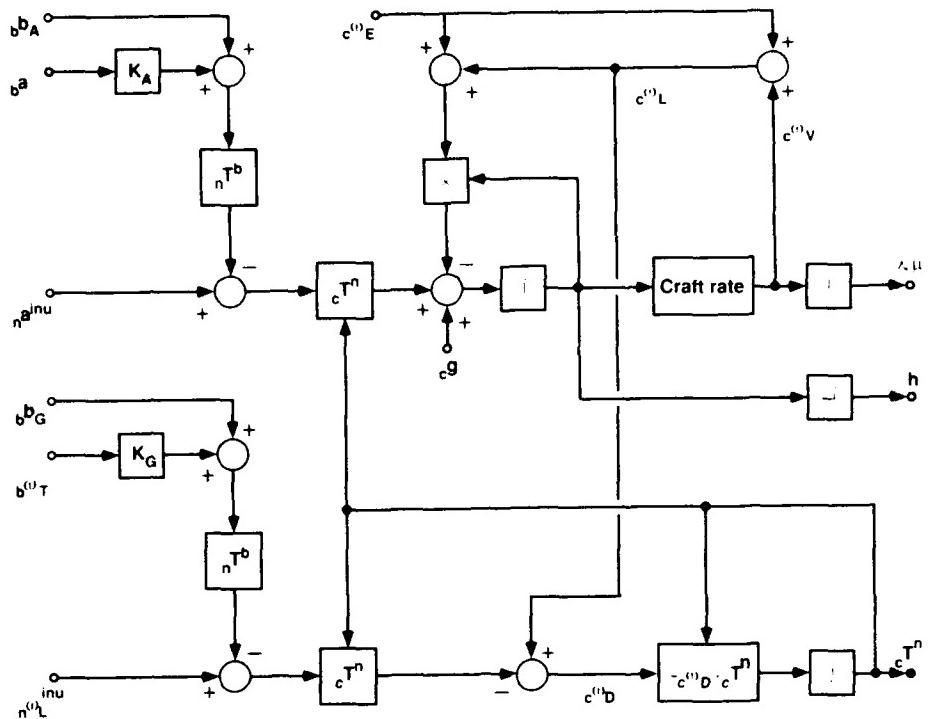


Fig. 4 - Inertial Navigation Equations

Note that the specific-force vectors $n^{inu} a^a$ and b_A^a , and the angular-rate vectors $n^{inu} \omega_L^a$ and b_T^a , and the rotation matrix $n^{inu} T^n$ in (7) and (9) are included with or can be computed from INU data.

Integration of GPS with INS using a range-domain Kalman filter is commonly used. One advantage to this technique is that the system will continue to navigate if range measurements from fewer than 4 satellites are available. Another advantage is that range measurements may be applied sequentially, thus the method is more compatible with sequencing receivers. However, the latter is a less significant advantage since receiver technology is moving towards multi-channel systems rather than sequencing systems.

The position-domain approach, which uses position and velocity residuals instead of range and range-rate residuals to drive the Kalman filter as indicated in Fig. 3., has received some attention in the literature [10,15]. In the position-domain solution, the DGPS solution was computed independently using a point solution [16] to the GPS equations with differentially corrected range data.

The point solution was extended to compute velocity using a least squares technique with the corrected range-rate data. One advantage to the position-domain filter is that the GPS clock bias and drift states are completely separate from the inertial navigation states. The clock correction states need not be included in the Kalman filter's state and covariance projection.

All of the DGPS/INS navigation systems discussed above (Figs. 1-3) integrate the navigation equations outside of the INU, i.e., in parallel with the INU's navigation solution. This was an experimental requirement. The DGPS/INS systems could be mechanized so that state and measurement errors are fed back into the INU if the unit would accept such inputs. The fully-integrated range-domain and position-domain navigation systems were implemented off-line and evaluated using GPS and INU data recorded during the flight-test. Appendix C describes, in greater detail, some of the design features of the actual off-line DGPS/INS implementations.

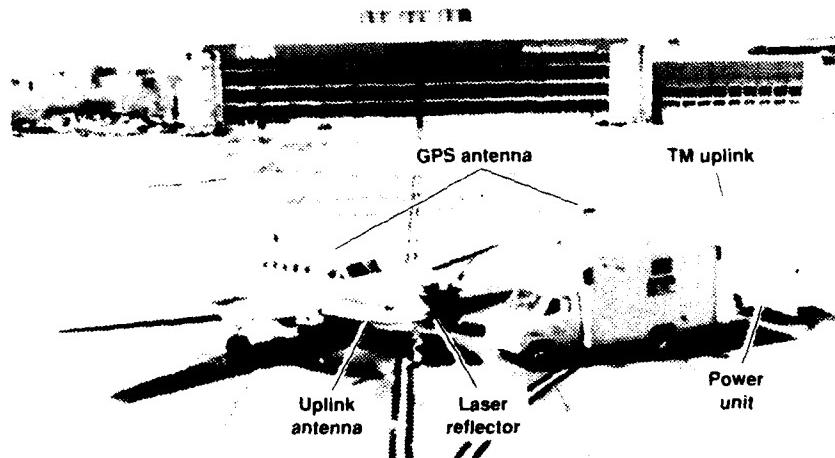


Fig. 5 - DGPS Test Aircraft and Ground Reference Station

Flight Test Data

The test aircraft (Beachcraft King Air 200, NASA 701) and the mobile DGPS ground reference station used during flight-test evaluation of P-Code DGPS/INS [5] are shown in Fig. 5. The airborne equipment included a 5 channel P-Code GPS receiver (Rockwell/Collins RCVR-3A), a strap-down ring-laser-gyro INU (Litton LN-93), and a real-time computer system. The ground system includes an identical GPS receiver and real-time computer system. Differential corrections were transmitted to the aircraft at 1 Hz using an L-Band telemetry uplink system.

Standard 3° approaches from 6.5 mi out were used to simulate approach and landing operations for fixed-wing aircraft. Flight tests were conducted at the Crows Landing Naval Air Station (located about 50 miles East of Moffett Field, CA) during the period from February to April 1991. Selective Availability was off during this time. The aircraft was laser tracked during each approach to determine its "true" position. Laser range accuracy is nominally ± 0.3 m (1σ noise) out to about 9 km; azimuth and elevation accuracy are nominally ± 0.2 milliradians (1σ noise). The laser tracker was calibrated in the morning before each day's flight test. Laser range, azimuth, and elevation data were zero-phase-shift filtered before being used in the truth position solution.

Since the RCVR-3A is effectively a 4 channel receiver, it was important that airborne and ground receivers tracked the same 4 satellites. The constraints associated with having only 4 channels at the reference station would normally be operationally unacceptable. An operational ground station should have multiple channels to allow tracking of all satellites in view. Differential corrections for all satellites should be uplinked to the aircraft. The airborne system should then choose satellites for its solution based on the availability of differential corrections from the ground system.

Approaches were selected from the data base for analysis based on three criteria: 1) continuous tracking of the same 4 satellites by the airborne and ground systems, 2) good satellite geometry, i.e., PDOP (position dilution of precision) < 6 , and 3) acceptable laser tracking data. Only approaches with good satellite geometry were used because when the GPS constellation is complete, good geometry will be available 24 hours a day from nearly every point on the globe. If the entire approach was not acceptable, the usable portion was included for analysis.

Results and Discussion

Positioning performance of the integrated navigation systems was evaluated based on statistical analysis of three axis (along-track,

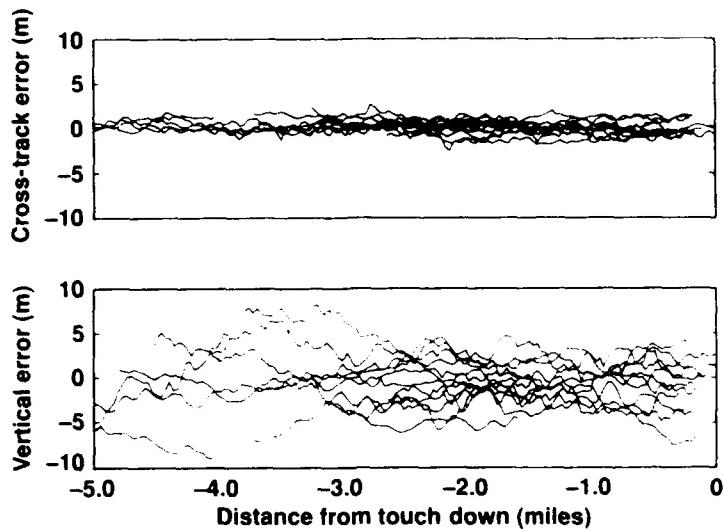


Fig. 6 - Cross-Track and Vertical Position Error Time Histories for 15-State Algorithm

cross-track, vertical) positioning errors during the final approach. Three axis position error time histories are computed by comparing the DGPS/INS navigation solution with the truth solution derived from laser tracker data. For the eight-state system, the real-time solution data recorded in flight were compared with the truth data. The fully integrated DGPS/INS range- and position-domain algorithms shown in Figs. 2 and 3 (also see Appendix C) were implemented off-line and validated using the flight-test data base. The off-line versions do not utilize the barometric altimeter so the range-domain filter has seventeen states instead of eighteen, and the position-domain filter has fifteen states instead of sixteen. Fifteen approaches flown on 4 different days with various satellite combinations were used in the post-flight analysis.

Figure 6 shows composite time histories of cross-track and vertical position errors for the 15 state position-domain solution plotted against true distance from the touch down point for all fifteen approach segments. As expected, the cross-track error is consistently smaller than the vertical error. Note that the vertical error bound tightens up at about 2.5 nmi (4.6 km), and then expands slightly during the last 1 nmi (1.8 km). This phenomenon is somewhat puzzling since GPS or DGPS accuracy should not be a function of position along the approach path. The fact that the laser

tracker is more accurate at shorter ranges could explain the better performance as the aircraft approaches the 2.5 nmi mark. The larger vertical error during the last mile could be due to increased multipath interference at the airborne antenna as the aircraft gets closer to the ground. However, this would only be a factor at altitudes below 30 m (0.3 nmi (0.5 km) on a 3° approach) since the P-Code wavelength is about 30 meters. The composite time history plots for the eight- and seventeen-state algorithms look similar to those shown in Fig. 6.

Figures 7-9 show cross-track and vertical position error histograms for the three levels of DGPS/INS integration. Three-axis position errors are separated into "error bins" according to axis and sign-magnitude. The histograms show the percentage of total points that fall into a particular error bin plotted against the sign-magnitude of that bin. The histograms are based on 1496 data points recorded at 1 Hz from the fifteen approaches.

Histograms provide a graphical representation of overall positioning performance for the fifteen approaches. The near-ideal Gaussian error characteristics of the histograms indicate that most correlated satellite ranging errors have been removed by the differential corrections. The along-track histograms (not shown) look nearly identical to the cross-track histograms, except that they

are offset by a bias error to be discussed later.

A Time History Analysis was performed which uses all data (at 1 Hz) from the error time histories in the statistical error analysis. This method is based on the assumption that DGPS positioning accuracy is not a function of aircraft position in the local area (e.g., position along the approach path). All points from each approach are given equal weight in the statistical analysis. Composite error statistics for the three levels of DGPS/INS integration are given in Table 1.

Decision point analyses are usually used to evaluate the navigational accuracy of approach and landing systems. To determine navigational accuracy, a decision point analysis uses only data points corresponding to a particular aircraft position along the glide path, e.g., 2 mi to touchdown (TD), or the 200 ft decision height (DH) on a 3 deg glide path. The mean and standard deviation of cross-track and vertical position error at 4 decision points along the 3° glide path were computed for the three levels of DGPS/INS integration. The results are summarized in Figures 10a (cross-track) and 10b (vertical). Though the DGPS mean and standard deviation appear roughly equivalent at each decision point along the glide path, there is likely some increase in positioning error at the 100 ft decision height as indicated in Fig. 6. The positioning accuracy of conventional approach and landing aids, such as the ILS or the Microwave Landing System (MLS), is inversely proportional to the distance to touch down. The positioning error statistics at the 200 ft. decision height (as plotted in Fig. 10) are given in Table 2.

The 200 ft decision height results in Table 2 indicate an apparent improvement in vertical axis performance for all systems as compared to the results of the Time History Analysis in Table 1. This reduction in error may be due to better performance when the aircraft is stabilized at the 200 ft decision height. However, these data are based on a limited data set, i.e., 11 approaches (11 data points).

The mean error in the along-track data is likely due to a time tagging error in either the airborne data or the laser tracking data, and not an error in the navigation solution due to along-track motion. The stand-alone GPS solution for the same data set [5] does not show an along-track mean error.

Therefore, we suspect that the along-track mean errors in Tables 1 and 2 can be calibrated to near zero. The vertical mean error is due to uncertainty in the DGPS vertical solution and can not be calibrated out of the solution. Therefore, it is appropriate to use the root-mean-square (RMS) error when measuring positioning performance.

During the course of this investigation many different tuning parameters, e.g., measurement noise and process noise, were used in the Kalman filters. Given the variation in results as a function of Kalman filter tuning, and the uncertainty in the laser tracking data, the variation in positioning error across the three levels of integration is considered statistically insignificant. This is not to say that higher levels of DGPS/INS integration are not beneficial. A major benefit of the full integration is to provide estimates of inertial instrument errors which may be applied to INS sensor data and, therefore, allow the INS to navigate more accurately should GPS data become unavailable. This benefit is not explored in this paper. Based on the results in Table 1, the average RMS positioning error across the three levels of integration is 1.0 m cross-track and 3.0 m vertical.

Another advantage of an integrated system is to provide smooth, calibrated, high rate navigation data between the 1 Hz DGPS measurement updates. Figure 11a is a sample approach time history showing the 20 Hz fifteen-state position-domain navigation solution and the DGPS measurement updates. Note that the 20 Hz navigation solution is a filtered fit to the DGPS data. The positioning error characteristics of the integrated solutions are dominated by the DGPS solution in both the range-domain and position-domain solutions. The Kalman filter adjusts the INS error estimates so that the 20 Hz navigation solution fits the DGPS data. Figure 11b is a short segment of the same approach. The quantization in the 20 Hz solution due to the DGPS updates is on the order of 10 cm. If necessary, (e.g., for guidance) this quantization could be filtered out independently of the DGPS/INS algorithm. The fits for the seventeen-state range-domain solution are similar to those in Figs. 11.

It was found that the use of deltarange measurements to approximate range-rate to the satellite introduced an error in the velocity estimation during turning flight. Since the

Table 1. P-Code DGPS/INS Results. Position Error Statistics (mean and 1-sigma standard deviation) from a Time History Analysis of data from 15 approaches (1496 data points) flown on 4 different days.

	8-State range-domain		17-State range-domain		15-State position-domain	
	mean	1σ	mean	1σ	mean	1σ
Cross-Track	0.1	0.9	0.3	1.0	0.2	1.0
Vertical	-0.9	3.0	-1.3	2.8	-1.4	2.4
Along-Track	-3.7	1.2	-3.3	1.1	-3.6	1.1

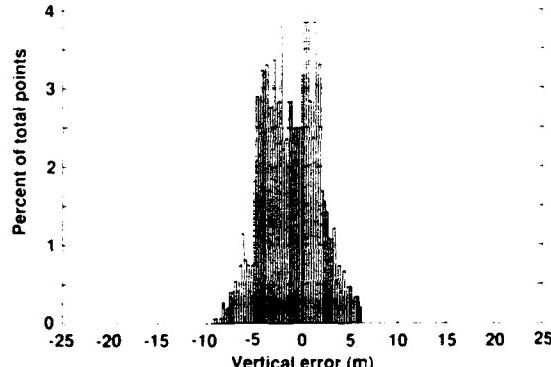
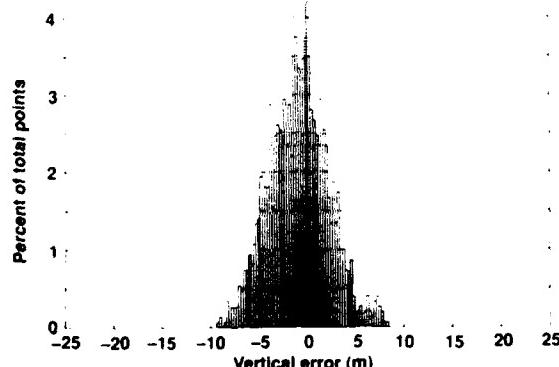
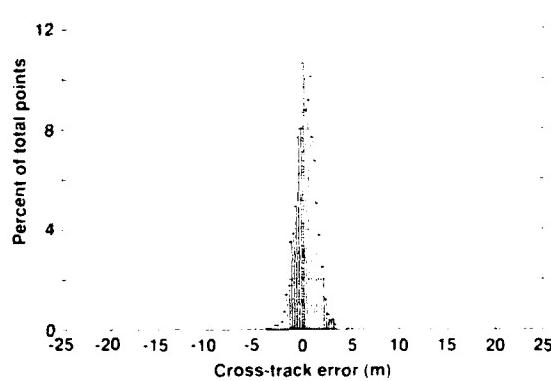
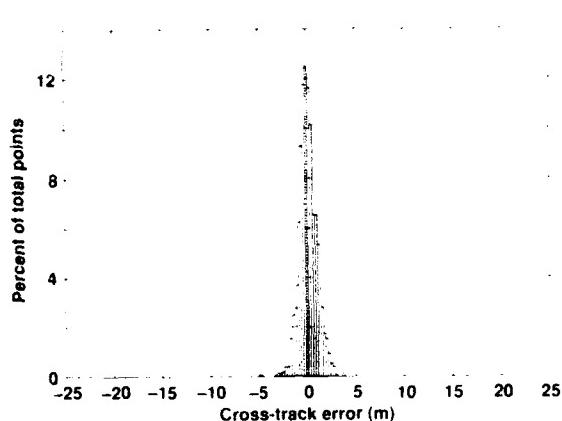


Fig. 7 - Histograms of Cross-Track and Vertical Position Error for 8-State Algorithm

Fig. 8 - Histograms of Cross-Track and Vertical Position Error for 17-State Algorithm

Table 2. P-Code DGPS/INS Results. Position Error Statistics (mean and 1-sigma standard deviation) at the 200 ft Decision Height based on data from 11 approaches (11 data points) flown on 4 different days.

	8-State range-domain mean 1 σ (meters)		17-State range-domain mean 1 σ (meters)		15-State position-domain mean 1 σ (meters)	
Cross-Track	-0.1	0.8	0.0	1.0	0.1	0.8
Vertical	-0.7	2.6	-0.8	2.2	-0.8	2.2
Along-Track	-3.5	1.1	-3.4	1.0	-3.5	1.3

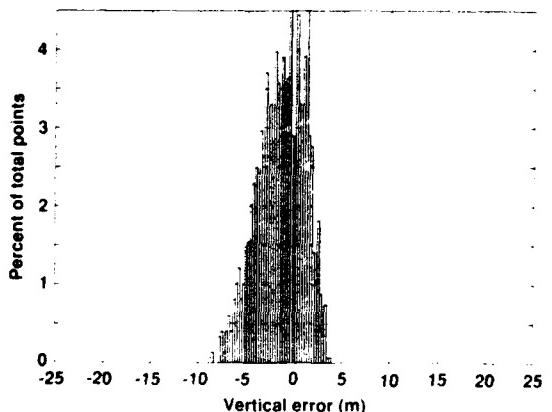
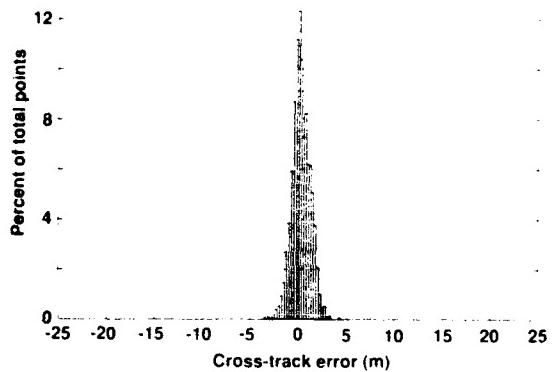


Fig. 9 - Histograms of Cross-Track and Vertical Position Error for 15-state Algorithm

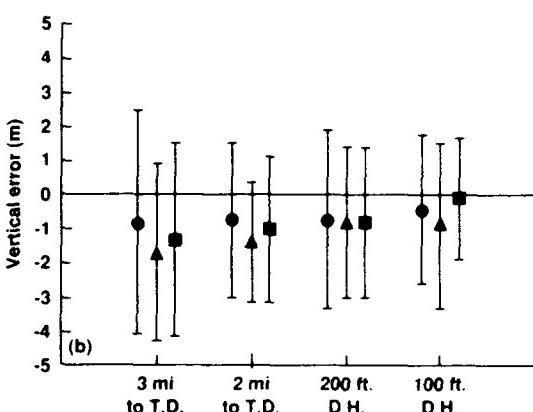
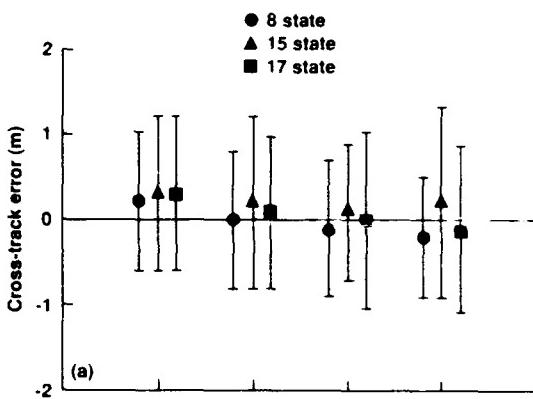


Fig. 10 - Mean and 1 σ Standard Deviation of Position Error at Decision Points on 3° Approach.

deltarange for the RCVR-3A is the range change during the 0.8 sec period prior to the pseudorange update, the range-rate estimate is the average range-rate over the 0.8 sec interval. At best, the range-rate estimate is valid 0.4 sec prior to the pseudorange-valid time. This latency was particularly apparent during turning flight. Only when the range-rate measurements were weighted low in comparison to the pseudorange measurements, would the Kalman filters converge properly during turning flight. The errors were not apparent during the steady approach since range-rate was relatively constant during this time. An improved method for making use of the deltarange measurements is currently being investigated.

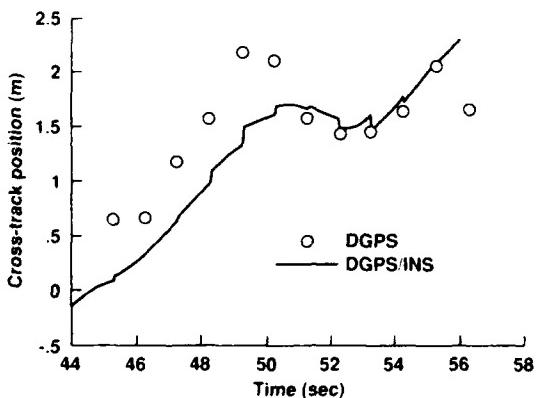
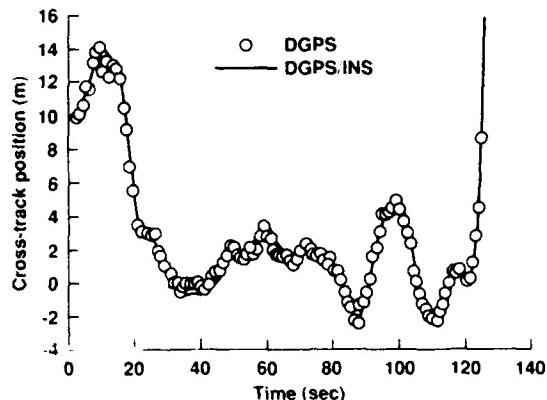


Fig. 11 - Cross-Track Position for Sample 3° Approach.

Concluding Remarks

Three different DGPS/INS integration algorithms were evaluated with flight-test data from a twin turbo-prop transport aircraft to determine positioning accuracy during approach and landing. A simple eight-state Kalman filter integration was implemented in real-time. Fully-integrated range-(seventeen-state) and position-domain (fifteen-state) algorithms were developed and evaluated off-line with the flight-test data base. Three-axis positioning accuracy (cross-track, vertical, along-track) during final approach was approximately the same for all three algorithms.

Though positioning performance is about the same for all levels of DGPS/INS integration, it is clear that the fully integrated systems which estimate INS instrument errors, i.e., accelerometer, rate-gyro, and barometric altimeter errors, would provide improved navigation accuracy during periods when GPS is unavailable.

DGPS/INS integration provided accurate, high frequency state estimates which may be used for navigation, guidance, and control. The eight-state system provided position and velocity estimates that are consistent with DGPS accuracy. The fully-integrated systems also provided calibrated estimates of acceleration, attitude, and body rate.

Average P-Code DGPS/INS positioning accuracy (root-mean-square) based on analysis of final approach flight-test data was 1.0 m cross-track and 3.0 m vertical. The DGPS/INS positioning accuracy results show about a factor of 3 improvement in both horizontal and vertical axis over that of stand-alone P-Code GPS, i.e., without differential corrections [5].

The use of GPS carrier phase data integrated with an inertial system is the subject of continuing research in GPS/INS integration at NASA Ames. GPS carrier phase measurements provide highly accurate range and range-rate information and could provide further improvements in navigation accuracy for terminal area flight operations.

Acknowledgment

The authors gratefully acknowledge the conversations held with Dr. Stanley F. Schmidt of Analytical Mechanics Associates concerning aided inertial

navigation systems.

References

1. Kremer, G.T., Kalafus, R.M., Loomis, P.V.W., Reynolds, J.C., The Effects of Selective Availability on Differential GPS Corrections, Navigation: Journal of the Institute of Navigation, Vol. 37, No. 1, Spring 1990.
2. Hogle, L., Investigation of the Potential Application of GPS for Precision Approaches, Navigation: Journal of the Institute of Navigation, Vol. 35, No. 3, Fall 1988.
3. Blackwell, E.G., Overview of Differential GPS Methods, Global Positioning System, Papers Published in Navigation, Vol III, The Institute of Navigation, 1986.
4. Milliken, R.J., Zoller, C.J., Principle of Operation of NAVSTAR and System Characteristics, Global Positioning System, Papers Published in Navigation, Vol I, The Institute of Navigation, 1980.
5. McNally, B.D., Warner, D.N., Hegarty, D.M., Schultz, T.A., Bronson, R., Flight Evaluation of Precision Code Differential GPS for Terminal Area Positioning, Proceedings of the 4th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-91), Albuquerque, New Mexico, September 11-13, 1991.
6. Browning, C.C., The NASA and Honeywell GPS/INS Flight Research Program, Presented at the First International Symposium Real Time Differential Applications of the Global Positioning System, Braunschweig, Federal Republic of Germany, Sept. 16-20, 1991.
7. Vallot, L., Snyder, S., Schipper, B., Parker, N., Spitzer, C., Design and Flight Test of a Differential GPS/Inertial Navigation System for Approach/Landing Guidance, Navigation: Journal of the Institute of Navigation, Vol. 38, No. 2, Summer 1991.
8. van Driel, N., Krijn, R., Results of a Test Program For the Use of Differential GPS For Approach Guidance, Proceedings of the Third International Technical Meeting of the Satellite Division of The Institute of Navigation, Colorado Springs, CO, Sept. 19-21, 1990.
9. Jacob, T., Dieroff, M., Integrated Navigation for Approach Guidance using Differential GPS, Proceedings of the 3rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-90), Colorado Springs, Colorado, Sept. 19-21, 1990.
10. Jacob, T., Schanzer, G., Integrated Flight Guidance System Using Differential-GPS for Landing Approach Guidance, AGARD Guidance and Control Panel, Lissabon, 1989.
11. Edwards, F.G., Loomis, P.V.W., Civil Helicopter Flight Operations Using Differential GPS, Navigation: Journal of the Institute of Navigation, Vol. 32, No. 3, Fall 1985.
12. Edwards, F.G., Paielli, R.A., and Hegarty, D.M., Helicopter Terminal Approach Using Differential GPS with Vertical-Axis Enhancement, Satellite Division Meeting of the Institute of Navigation, Colorado Springs, CO, Sept. 1987.
13. Edwards, F.G., D.M. Hegarty, R.N. Turner, F. van Grass, and S. Sharma, Validating the Airborne and Ground Based Components of a Differential GPS System, National Technical Meeting of The Institute of Navigation, Santa Barbara, CA, Jan. 1988.
14. Edwards, F.G., Hegarty, D.M., Flight Test Evaluation of Civil Helicopter Terminal Approach Operations Using Differential GPS, AIAA Guidance, Navigation, and Control Conference, August 14-16, 1989. AIAA 89-3635.
15. Chaffee, J.W., Abel, J.S., The GPS Filtering Problem, IEEE PLANS '92 Symposium, Monterey, California, March 23-27, 1992.
16. Bancroft, S., An Algebraic Solution of the GPS Equations, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES 21, No. 7, January 1985.

Appendix A - Pseudorange, Deltarange, and Differential Correction Processing

The DGPS reference station computes corrections to pseudorange and range-rate data for all satellites being tracked by the ground receiver. Differential corrections are filtered at the ground

station and data-linked to the airborne system at a rate of 1 set per second; a set includes range and range rate corrections for all satellites. The airborne system applies the differential corrections to corresponding satellite measurements from the airborne receiver.

The receiver's raw pseudorange and deltarange measurements, its range error estimates, the satellite ephemeris, and its GPS solution data, including estimates of clock bias and clock drift, are all output from the receiver. Pseudorange measurements in both the airborne and ground reference systems are first adjusted with the receiver's estimate of total range error, less the ionospheric contribution which will be accounted for by the differential corrections. Range-rate is approximated by dividing the deltarange measurement by the time interval over which the deltarange was valid. Pseudorange and range-rate are then adjusted with the receiver's estimate of clock bias and clock drift. This removes most of the receiver's clock error from the pseudorange and range-rate measurements. The airborne navigation filter still estimates any residual clock bias and clock drift. The equations for computing range and range-rate estimates in both the airborne and ground systems are

$$\rho = \rho - e_r - b_r - d_r (\Delta t_{drift}) \quad (A.1)$$

$$\dot{\rho} = \Delta r / \Delta t_{int} - d_r \quad (A.2)$$

At the ground reference station differential corrections are computed by differencing the range and range-rate estimates from (A.1) and (A.2) with the computed range and range rate as determined using satellite position and velocity and the surveyed reference station coordinates. The differential range and range-rate corrections are given by

$$\Delta\rho = \rho_0 - \rho \quad (A.3)$$

$$\Delta\dot{\rho} = \dot{\rho}_0 - \dot{\rho} \quad (A.4)$$

These data are filtered at the ground station, uplinked to the aircraft, and applied to corresponding airborne satellite measurements.

In the airborne system, corrected range and corrected range-rate are given by

$$R = \rho + \Delta\rho + \dot{\rho} (\Delta t_{lat}) \quad (A.5)$$

$$\dot{R} = \dot{\rho} + \Delta\dot{\rho} \quad (A.6)$$

The uplink latency, Δt_{lat} , is the difference between the GPS time tag on the airborne measurements and the GPS time tag on the differential correction message.

Appendix B - Derivation of the Tilt Matrix Kinematics

The integrated navigation system described in this paper operates in parallel with a standard inertial navigation unit as shown in Figs. (1) and (2). The inputs to the navigation equations consist of specific-force measurements in the INU level frame, and body frame angular rate-gyro measurements. This Appendix presents a derivation for the kinematics of the "tilt" matrix required to transform a vector from the INU level frame ("n") to the corrected level frame ("c"). All matrices are (3x3), and vectors are (3x1). The derivation of the tilt matrix kinematics begins with the rotation-matrix identity

$${}^cT^n = {}^cT^b {}_bT^n \quad (B.1)$$

Now, differentiate both sides of (B.1) to obtain

$$\dot{{}^cT^n} = {}^c\dot{T}^b {}_bT^n + {}^cT^b \dot{{}_bT^n} \quad (B.2)$$

and use the matrix identity

$$\dot{{}^cT^b} = - {}^cT^b {}_b\dot{T}^c {}^cT^b \quad (B.3)$$

to write (A.2) as

$$\dot{{}^cT^n} = - {}^cT^b ({}_b\dot{T}^c {}^cT^n - {}_b\dot{T}^n) \quad (B.4)$$

The kinematics for the level-to-body frame transformation calculated by the INU are given by

$$\dot{{}_bT^n} = - ({}_b\omega_T - {}_bT^n {}_n\omega_L^{inu}) \times {}_bT^n \quad (B.5)$$

where

$${}_n\omega_L^{inu} = {}_n\omega_V^{inu} + {}_n\omega_E \quad (B.6)$$

The aided navigation system should be able to estimate any residual errors in the total angular rates not compensated by the INU. These errors can be modeled in the corrected level frame as

$$\begin{aligned} \dot{\mathbf{b}}^T \mathbf{c} &= -(\mathbf{b}\omega_T - \mathbf{K}_G \mathbf{b}\omega_T - \mathbf{b}b_G \\ &\quad - \mathbf{b}^T \mathbf{c} \omega_L) \times \mathbf{b}^T \mathbf{c} \end{aligned} \quad (\text{B.7})$$

where \mathbf{K}_G is an upper triangular matrix of rate-gyro scale and misalignment factors, and $\mathbf{b}b_G$ is a vector of rate-gyro biases, in the body frame, and where

$$\mathbf{c}\omega_L = \mathbf{c}\omega_V + \mathbf{c}\omega_E \quad (\text{B.8})$$

Finally, with substitution of (B.5) and (B.7) into (B.4) and with the rotation matrix identity $\mathbf{T}(a \times) = (\mathbf{T}a)\mathbf{x}\mathbf{T}$, the kinematics required for calculating the tilt matrix can be written in the form

$$\begin{aligned} \dot{\mathbf{c}}^T \mathbf{n} &= -\{\mathbf{c}\omega_L - \mathbf{c}^T \mathbf{n} [\mathbf{n}\omega_L]^{\text{inu}} \\ &\quad - \mathbf{n}^T \mathbf{T}^b (\mathbf{K}_G \mathbf{b}\omega_T + \mathbf{b}b_G)\} \times \mathbf{c}^T \mathbf{n} \end{aligned} \quad (\text{B.9})$$

Note that the angular-rate vectors $\mathbf{n}\omega_L^{\text{inu}}$ and $\mathbf{b}\omega_T$, and the rotation matrix \mathbf{T}^b in (B.9) can be computed from INU data.

Appendix C - Post Flight DGPS-Aided Navigation Filter

In the flight-test experiment described in this paper, GPS data at 1 Hz and INU data sampled at a nominal rate of 20 Hz were stored on disk. These data were used in the post-flight evaluation of the DGPS-aided navigation systems. The INU-derived measurements of specific forces and angular rates were used to drive the parallel inertial-navigation computations outlined in Fig. 4. Notice that body-frame accelerometer and rate-gyro errors are each modelled with a bias vector and a scale-factor (misalignment) matrix. The purpose of the navigation filter is to provide an "in-flight" calibration (or identification) of the error models and thereby provide precise navigation when DGPS data are available. This appendix describes some of the design features of the post-flight DGPS-aided navigator shown in Fig. C.1.

It was decided that a wander-angle mechanization would add unnecessary complexity to the navigation problem. The coordinate frame chosen for the parallel navigator was a North, East, down locally-level system common in aeronautics. As indicated earlier, the navigation states comprised INU position (geodetic latitude, longitude, altitude) and

velocity (North, East, down), and three tilt angles (to rotate from the INU level to the corrected level frame). Both the position- and range-domain versions of the post-flight filter utilize the same navigation equations ((6)-(9) in the text). The navigation states are integrated at the nominal 20 Hz rate, using a second-order Runge-Kutta method.

The integration time step is variable, and is determined as the difference between consecutive INU times:

$$\Delta t_{\text{INU}} = t_{\text{INU}}^2 - t_{\text{INU}}^1 \quad (\text{C.1})$$

In the navigator actually implemented, the scale-factor matrices \mathbf{K}_A and \mathbf{K}_G were not included. The individual test maneuvers were only 3-5 minutes in duration and quite benign, so that little benefit would accrue beyond estimating accelerometer and rate-gyro biases. Therefore, the state model for the position-domain filter has fifteen states, while the state model for the range-domain filter, which must estimate clock bias and clock drift, has seventeen states. The corresponding filter measurement models have dimension six (position and velocity), and eight (range and range rate to four satellites), respectively. It should be noted that the DGPS measurements were referenced to the antenna location, requiring "lever-arm" corrections to position and velocity for calculation of measurement residuals.

It should be noted that in this filter implementation, the process noise enters the state only through the bias models. Each model is of the form

$$\dot{\mathbf{b}} = -\mathbf{b}/T + \mathbf{w} \quad (\text{C.2})$$

This is recognized as a "colored-noise" model, with time constant T and zero-mean white-noise source w. A convenient discrete-time form of this bias model is given by

$$\begin{aligned} \mathbf{b}(i+1) &= (1 - h/T) \mathbf{b}(i) \\ &\quad + (2h/T)^{1/2} \mathbf{w}(i) \end{aligned} \quad (\text{C.3})$$

where h is the filter time-update interval (Δt_{INU} in Fig. C.1), and the noise sequence $\mathbf{w}(i)$ has a variance equal to that of the apriori estimate of the bias error itself, i.e.

$$E\{\mathbf{w}^2(i)\} = \sigma_b^2 \quad (\text{C.4})$$

The model has the property that, in the absence of measurement updates, the steady-state value of the bias variance is exactly σ_b^2 .

Some comments concerning calculation of the state transition matrix are in order. As indicated in the text, the navigation state model is nonlinear, and can be represented in vector form as

$$\dot{x} = f(x, u, w), \quad x(0) = x_0 \quad (C 5)$$

where x is the navigation state vector (including biases), u is the forcing-function vector (INU specific forces and angular rates), and w is a representation of the modelling error (process noise). The state-error equation for the Kalman filter is of the form

$$\delta\dot{x} = f_x \delta x + f_w \delta w \quad (C 6)$$

where f_x and f_w are the partial-derivative matrices. Note that the state partial f_x is a function of the navigation states and must be recalculated at each point of the state trajectory.

For this filter implementation, as in most applications, it is sufficient to approximate the discrete-time transition matrix by

$$\Phi = [I + F + F^2 / 2], \quad F = \sum f_x dt \quad (C 7)$$

where the term F is the result of accumulating products $f_x dt$ computed at each time step (deltN) of the navigation solution during the interval between filter time updates (deltU). Hence, when a time update is to be performed, most of the work required to compute Φ has already been done.

Because the navigation sampling times were unsynchronized (nominally 20 Hz), the Kalman filter was designed to run synchronously with the DGPS measurement schedule, which was periodic at 1 second. Although the filter time-update interval can be a sub-multiple of this period, the results presented in the text were all obtained with one-Hz time updates. Refer to Fig. C 1 and note that, in general, a time update at ttup or a measurement update at tgps would occur between tins¹ and tins². Before performing the time update, the term F in (C 7) must be completed for the interval

$$delt1 = ttup - tins^1 \quad (C 8)$$

Following the time update, F must be initialized for the interval

$$delt2 = tins^2 - ttup \quad (C 9)$$

In order to perform the measurement update, the navigation state must be interpolated between tins¹ and tins² in order to calculate the measurement residual at tgps. Furthermore, following the update, the state error must be extrapolated to tins² in order to be applied to the navigation state. Finally, notice that the state error is reset to zero before exiting the measurement-update task.

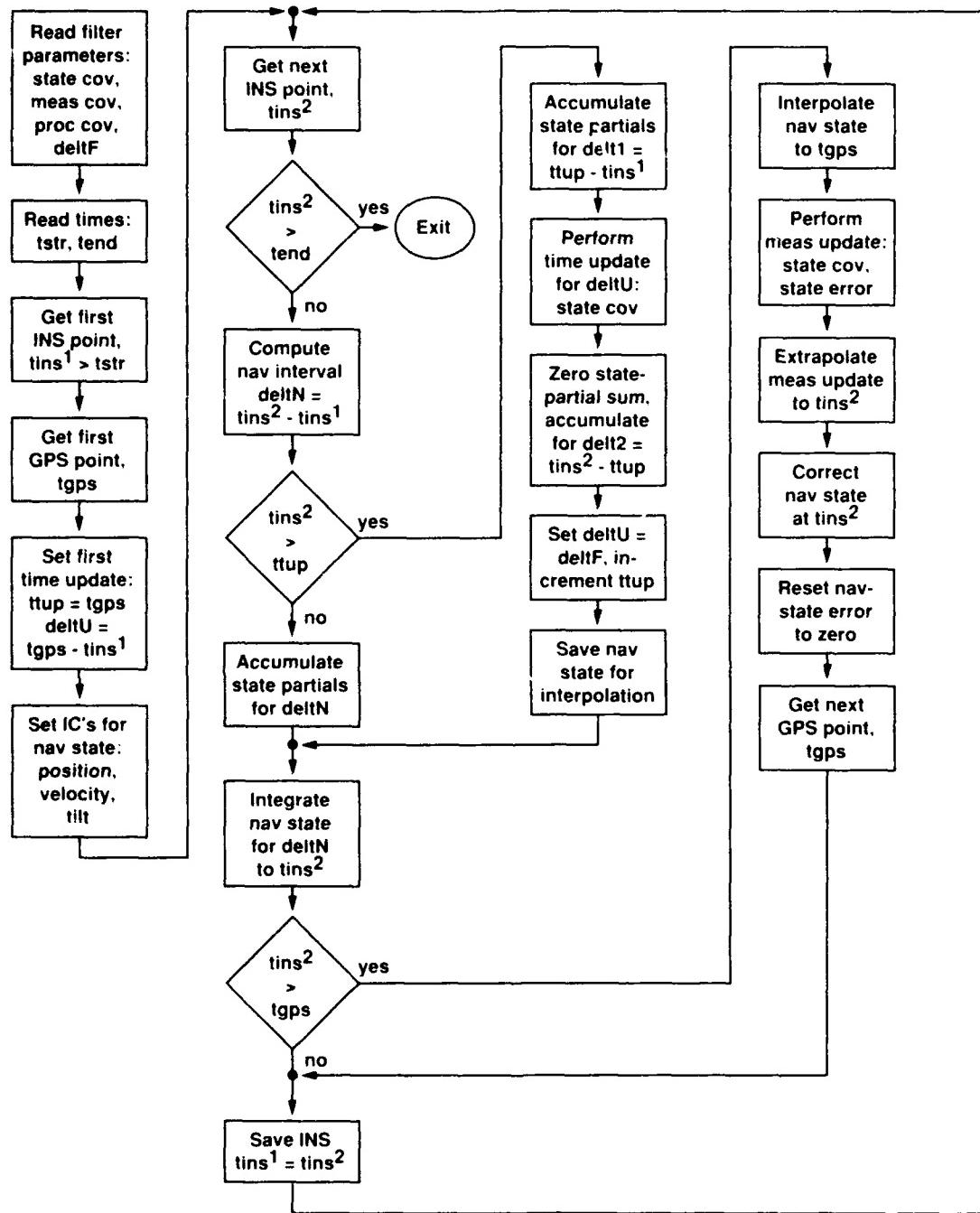


Fig. C1 - Post-Flight DGPS-Aided Navigation Filter.

Visual Autonomous Automatic Landing of Airplanes

E.D. Dickmanns, F.-R. Schell

Aerospace Technology Department
 Universität der Bundeswehr München
 W. Heisenberg Weg 39
 D-8014 Neubiberg
 Federal Republic of Germany

Summary

A visual sensor data processing method has been developed and validated which allows to achieve on-board autonomous landing approaches in the visual flight regime with computing technology available today; sensors are a video-camera, inertial gyros and an air velocity meter. The key feature of the method is the reconstruction and servo-maintained adjustment by prediction error feedback of an internal spatio-temporal model about the process to be controlled (4D approach). This encompasses both the egomotion state of the aircraft carrying the sensors and the relevant geometric properties of the runway and its spatial environment. The efficiency of the approach is proved both in a hardware-in-the-loop simulation and in real test-flights with a twin-turbo-prop aircraft Do 128 of Dornier. For accuracy evaluation of the data gathered, the results of differential GPS and radiometric altitude measurements have been recorded simultaneously.

List of Symbols

ADC	analog digital converter
BVV	image sequence processing system
DAC	digital analog converter
DBS	three-axis motion simulator
u, v, w	translational velocities of the airplane
p, q, r	angular velocities
Φ, Θ, Ψ	Euler angles
x, y, H	distance of the airplane to the runway
η	elevator angle
F	thrust
ξ	aileron angle
ζ	rudder angle
x	state vector
u	control vector
Φ_{LS}	transition matrix (longitudinal, lateral motion)
A	Jacobian of system model
H	Jacobian of measurement model
K	Kalman gain matrix
P	error covariance matrix
Q	covariance matrix of system noise
R	covariance matrix of measurement noise
G	input matrix of noise process

1. Introduction

Electronic micro-miniaturization of sensors and processors is progressing to a stage where machines may be provided with the equivalent of the human sense of vision. Only a few years ago, the 1 million-instructions-per-second (MIPS) performance class for digital computers has been a magic limit; within a few years the 'GIPS'-class (Giga, i.e. 10^9 instructions per second) will be commonplace. This will allow to process high data rates as produced by imaging sensors in real time. Color video requires a data rate of the order of magnitude of 10 MB/s.

However, data rate is not the essential point since it is the information content of an image which is useful for achieving some goal based on image sequence processing. Within a high frequency image sequence there may be quite a bit of redundancy since the situation changes only slowly over time, in general. Therefore, the main task of real-time image sequence processing is to reduce data rates but to keep as much information about the process to be controlled as possible.

A uniformly grey image contains as many picture elements (that means 8 bit data points) as a highly structured one; yet, the information content of the former may be summarized completely (*without any loss*) by 1. the symbol 'uniformly grey' and 2. the number coding the grey level. For a $1K \times 1K$ pixel image this corresponds to a data reduction of the order of 10^6 .

This is well appreciated in static image processing where segmentation of regions with similar characteristics is a generally accepted first step; regions or contour models allow much denser representation and storage of information than handling individual pixels. However, the same has not been true along the temporal axis in most approaches to image sequence processing. The 4D approach developed at UniBWM [1 to 5] combines both spatial and temporal models about processes in the world and fully exploits continuity conditions along all 3D space axes and along the time axis simultaneously, hence the name '4D approach'.

In this approach, all processing activities are geared to the next point in time when new measurements are going to be taken. There is no storing of previous measurement data for differencing or rate computation; this is of especial interest in image sequence processing where each

measurement means huge amounts of data (10^5 to 10^6 Bytes), however, very much less new information once the notion of objects and their states has been introduced. The results of previous measurements and evaluations are stored in parameters and state variables of generically (structurally) defined object models including their motion behavior. In modern control theory this procedure is well known as recursive estimation (Kalman filters, Luenberger observers). This has been extended to perspective mapped image sequences and was shown to be numerically very efficient. The flexibility of the approach has been demonstrated in the application areas of road vehicle guidance [2, 3], satellite docking [4], landmark navigation for autonomously guided vehicles on the factory floor and for landing approaches of aircraft. The latter one is the most demanding application up to now and will be discussed in the sequel.

2. Multi-point model of airplane dynamics

The most pretentious application of the 4-D approach is the automatically controlled landing approach of an airplane, because here a body is able to move within all six degrees of freedom (three translational and three rotational). According to Newton's law a state vector with 12 components is necessary for the description of the complete dynamics of the airplane. These equations are nonlinear. Contrary to the often used one-point airplane models, for this application a multi-point model description is used. The aerodynamic forces and moments are modelled separately on the wing and on the elevator unit (fig. 1) [6]. Wind effects, which have a considerable influence on the aircraft dynamics are included in this model too.

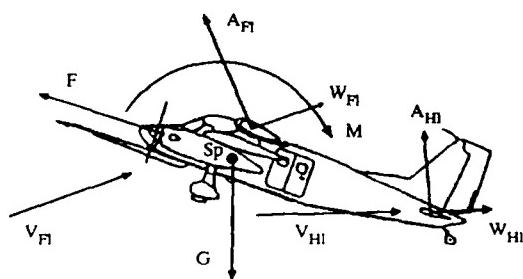


Fig. 1: Multi-point model of the Domier Do 128 airplane

The state vector \mathbf{x} consists of the translational (u, v, w) and angular (p, q, r) velocities, the Euler angles Φ, Θ, Ψ , and the distance to the middle of the runway threshold (x, y, H), which is the origin of the coordinate system chosen. The four component control input vector \mathbf{u} is assembled of the elevator angle η , thrust F , the aileron angle ξ and the rudder angle ζ . With this, the set of first order nonlinear differential eqs. may be written in the standard form for a 'dynamical model':

$$\mathbf{x} = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), (t)] \quad (1)$$

After linearisation around a (sliding) reference point $\mathbf{x}_0, \mathbf{u}_0$, the 12-th order system splits into two loosely coupled 6-th order systems: the longitudinal one with

$$\mathbf{x}_L = (u, w, q, \Theta, x, H)^T; \mathbf{u}_L = (\eta, F)^T \quad (2a)$$

and

$$\dot{\mathbf{x}}_L = \mathbf{A}_L[\mathbf{x}_0, \mathbf{u}_0] \mathbf{x}_L + \mathbf{B}_L[\mathbf{x}_0, \mathbf{u}_0] \mathbf{u}_L; \quad (2b)$$

the lateral one with

$$\mathbf{x}_S = (v, p, r, \Phi, \Psi, y)^T; \mathbf{u}_S = (\xi, \zeta)^T \quad (3a)$$

and

$$\dot{\mathbf{x}}_S = \mathbf{A}_S[\mathbf{x}_0, \mathbf{u}_0] \mathbf{x}_S + \mathbf{B}_S[\mathbf{x}_0, \mathbf{u}_0] \mathbf{u}_S. \quad (3b)$$

The linear systems are the basis for developing a feedback controller, while motion simulation is performed using the original nonlinear equations (1).

3. Visual measurement model

For the imaging process from radiating points in 3D space onto the image plane the simple pinhole camera model is adopted (straight line perspective mapping). A point P in the runway plane has the coordinates (x_L, y_L, z_L) in an axis frame with the origin in the center of the runway threshold and the x-axis aligned with the runway center line (fig. 2). The position of the airplane in this coordinate system is at point (x, y, z) , where a geodetic coordinate system with z_g in the direction of the Earth gravity vector is affixed to the aircraft center of gravity (cg); the x_g -axis in the horizontal plane is usually defined towards geographic north. For a right handed system the y_g axis then points towards east.

The angular orientation of the aircraft relative to this geodetic system is given by the three Euler angles Ψ, Θ, Φ , where the sequence of rotation is of importance for the final orientation; here, the z-sequence (Φ, Θ, Ψ) has been used since it yields relatively simple results in combination with the viewing direction control. In the aircraft-oriented coordinate system indexed f , the projection center of the camera has the coordinates (l_x, l_y, l_z) . This is the origin for the camera-oriented coordinate system (indexed k), the angular orientation of which relative to the f -frame is ψ_{fa} around the z_f -axis and θ_{fa} normal to the xy_f -plane, positive upwards. The point P is mapped into the image plane at distance f (focal length of camera lens) normal to the x_{fa} -direction with the coordinates z_{By} in line-direction (horizontally) and z_{Bz} in column-direction (vertically).

Using homogeneous coordinates, the transformations can be easily computed by 4*4 matrix multiplications; the following sequence is applied: From runway-coordinates translation T_g into geodetic coordinates, rotation R_f into airplane coordinates, translation T_c into platform-base coordinates, rotation R_{Ca} into camera coordinates and perspective projection P into image coordinates (for details see [9]). The nonlinear overall mapping equation may be written in vector form for the two image coordinates \mathbf{z} , with \mathbf{p} as camera mapping parameter vector

$$\mathbf{z} = \mathbf{h}[\mathbf{x}, \mathbf{p}] \quad (4)$$

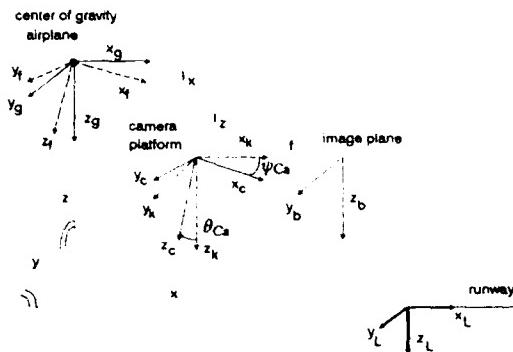


Fig. 2: Mapping a point from the runway into the image plane of the camera [9]

With $dx = X_L - x$, $dy = Y_L - y$, $H_z = \text{height above ground } (-z)$, and the denominator

$$D = b_{41} dx + b_{42} dy + b_{43} H_z + b_{44}, \quad (4a)$$

a point in the runway plane at $(X_L, Y_L, 0)$ will be mapped into the image plane at

$$z_{By} = (b_{21}dx + b_{22}dy + b_{23}H_z + b_{24})/D \quad (4b)$$

$$z_{Bz} = (b_{31}dx + b_{32}dy + b_{33}H_z + b_{34})/D \quad (4c)$$

where the b_{ij} are coefficients depending on the transformation parameters (see [9]). Real measurements are always noise corrupted; therefore, for image interpretation through recursive estimation an additive noise term $v(t)$ is assumed to be present with covariance matrix R . The Jacobian matrix of the right hand side of eq. (4) taken with respect to the aircraft state x is abbreviated with H (see eq. (7) below); this matrix and the coefficients b_{ij} become especially simple if the viewing direction is fixation controlled towards a point at the horizon where the runway borderlines intersect each other. This is obtained by a two-axis platform on which the camera is mounted. This platform is able to move in azimuth and elevation thus trying to keep the picture of the runway in the center of the image plane.

Eq. (4) is evaluated at ten different points in the runway plane where linearly extended intensity gradients may easily be found by intelligently controlled correlation with gradient templates (elongated ternary masks), see figure 3.

Knowing the shape of the runway, usually a rectangle, the appearance of the borderlines under perspective projection can be computed from the four corner points, given the relativ aspect conditions - in the definition chosen, exactly the aircraft state components. Eight windows are placed on the runway boundaries and two on the horizon in order to determine the roll angle.

Once an initialisation has been achieved, the search regions within the windows can be kept small since, due to motion prediction exploiting the dynamical model and previous control inputs, only the effects of disturbances have to be compensated by the search. Systematic changes in perspective projection are taken into account since all internal representations are simultaneously in 3D

space and time. This is equivalent to what psychologists call the 'Gestalt' phenomenon: When it is known what to look for, the interpretation of a scene may be much easier and less ambiguous than without any previous knowledge. This has been of great help in road recognition when, due to shadows from trees, intensity gradients are abundant and the highest correlation values do not at all correspond to road boundaries. Runway recognition, usually, is much more simple; however, taxiway entries and exits may be compensated for by dropping the measurements in these areas.

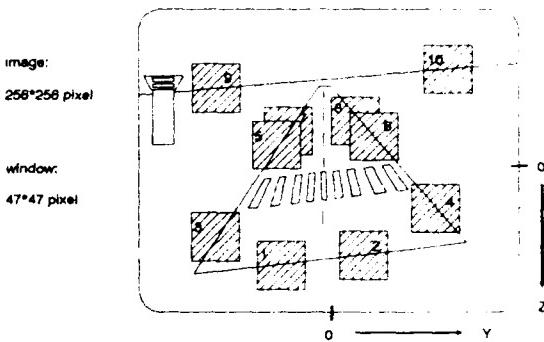


Fig. 3: Regions of evaluation (windows) within one image

4. Filter model for estimation of the complete state vector in real-time

In order to establish an Extended Kalman Filter (EKF) for the estimation of the complete state vector in real-time, the nonlinear equations of motion (1) are used. Based upon this 'dynamical model' the fundamental algorithm of an EKF may be found in [8] and looks like follows:

The transition matrix Φ over one cycle period $\Delta t = t_{i+1} - t_i$ is obtained from the linearised nonlinear system equations (1) with the Jacobian A:

$$A[t; \hat{x}(t | t_i)] = \frac{\partial f[x(t), u(t), t]}{\partial x} \Bigg|_{x = \hat{x}(t | t_i)} \quad (5)$$

Φ is defined as

$$\Phi[t_{i+1}, t_i; \hat{x}(t | t_i)] = e^{\Phi(t_i) \Delta t} \quad (6)$$

P denotes the error covariance matrix, Q and R the covariance matrices of the system and measurement noises, while G is the input matrix of the noise process.

H is the Jacobian of the right hand side of the nonlinear measurement equations $h[x, p, t_i]$ with p as parameters:

$$H[t_i; \hat{x}(t_i)] = \frac{\partial h[x, p, t_i]}{\partial x} \Bigg|_{x = \hat{x}(t_i)} \quad (7)$$

The final algorithm can be written as:

Extrapolation

$$\hat{x}(t_{i+1}^-) = \hat{x}(t_i^+) + \int_{t_i}^{t_{i+1}} f[\hat{x}(t, t_i), u(t, t)] dt \quad (8)$$

$$P(t_{i+1}^-) = \Phi P(t_i^+) \Phi^T + \int_{t_i}^{t_{i+1}} \Phi G(t) Q(t) G^T(t) \Phi^T dt \quad (9)$$

Innovation (with $H = H[t_i; \hat{x}(t_i^-)]$)

$$K(t_i) = P(t_i^-) H^T (H P(t_i^-) H^T + R(t_i))^{-1} \quad (10)$$

$$\hat{x}(t_i^+) = \hat{x}(t_i^-) + K(t_i) [z_i - h[\hat{x}(t_i^-), t_i]] \quad (11)$$

$$P(t_i^+) = P(t_i^-) + K(t_i) H [t_i; \hat{x}(t_i^-)] P(t_i^-) \quad (12)$$

In the extrapolation step (eq. 8 and 9) the system and corresponding error models are integrated from the actual point in time (t_i^+) to the next one (t_{i+1}^-). After having received the new measurement values, the innovation is performed (eqs. 10 to 12), yielding the best estimate for the system state at time (t_i^+), which is in turn the basis for the next filter step.

Up to now, low cost microprocessor hardware is too slow for implementing the complete algorithm on a single unit in real-time. In order to reach real-time performance despite these difficulties, some steps have been investigated in order to split the algorithm from one set for 12 vector components into 2 sets indexed L and S for 6 vector components each running on parallel processors. Since the amount of computation needed is proportional to n^3 , this reduces the computation time needed to 12.5 %. Finally, real-time performance has been achieved by substituting in the separated equation (9) the following terms:

$$\int_{t_i}^{t_{i+1}} \Phi_L Q_L(t) \Phi_L^T dt = \bar{Q}_L \quad (13)$$

$$\int_{t_i}^{t_{i+1}} \Phi_S Q_S(t) \Phi_S^T dt = \bar{Q}_S \quad (14)$$

This allows to also split eqs. 9 to 12 into two sets resulting in two decoupled systems for the aircraft motion (fig. 4).

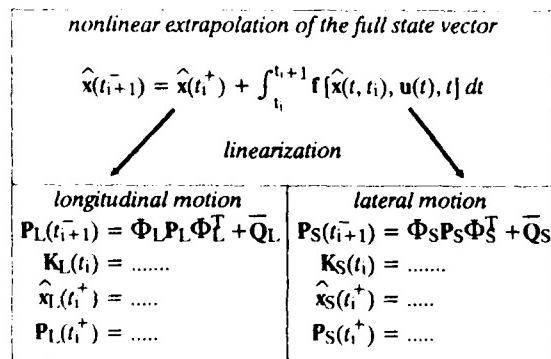


Fig. 4: Splitting the filter algorithm into two separated parts

These separate main parts have been implemented on two different processors and run in parallel. For additional speedup, the sequential algorithm of Bierman has been used [12]. With this, the crucial step to real-time with $\Delta t = 60 ms$ cycle time was achieved. By sticking to the nonlinear extrapolation for the system state, no knowledge about the system is lost.

The bulk of the measurement vector consists of optical features which are extracted from a single image delivered by a video camera. The rest of the measurements are inertial values obtained by gyros fixed to the aircraft. As indicated in section 3, a feature is a linearly extended intensity gradient within the runway plane. It is extracted by controlled correlation with ternary masks on the runway borderlines and on the horizon along special search paths [3]. The feature extraction algorithms are implemented on different parallel processors PPi in the image processing system BVV (see fig. 5), where one processor analyzes two features within 40 ms. In this approach maximally ten features are used per cycle presently (fig. 3).

5. Hardware

The main part of the underlying hardware is the image processing system BVV, developed at the Universität der Bundeswehr München [5, 7]. The BVV is a multi-processor system in which the different processors can communicate via a common system bus. Each pixel processor PPi (Intel 80286) has access to the video image from a CCD camera via a video bus and an ADC (fig. 5).

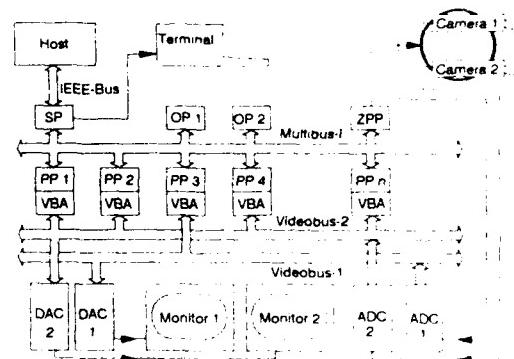


Fig. 5: Multi-processor system BVV (from [3])

The object processors OPi (80386) run the recursive estimation algorithms based on the 'Gestalt' idea of a perspectively mapped runway and the dynamical models, exploiting the feature data in conjunction.

Integrated into the BVV is the controller for the two-axis platform (ZPP). The host for the image processing system is a PC; both are connected by an IEC-bus. The PC is used only for initialization, as a link to the integrated computer of the airplane, and for data collection and final evaluation. For visual control of the process, a video monitor and a video recorder have been integrated (fig. 6).

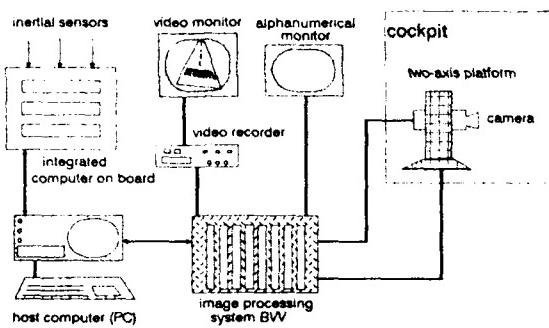


Fig. 6: Hardware architecture

6. Hardware-in-the-loop (HIL) simulation

Primary results have been achieved in a hardware-in-the-loop simulation, which was especially built for investigations and developments in the field of image processing [10]. Real sensing and processing components may be integrated into this loop. The dynamics of the airplane and its view onto the runway can be simulated by a three-axis rotational motion simulator (DBS) and a graphics system (SGI 4D). The integration of the non-linear equations of motion is done by a digital computer which controls the other simulation facilities too (fig. 7).

For the simulations within the entire flight envelope, a complete state feedback controller has been developed with the latest theory of linear quadratic design with prescribed eigenstructures [11]. With this method, the Riccati design is combined with the pole placement and eigenvector specification, thus allowing to use the advantages of both methods. The nominal trajectory to be flown in 3D-space and time is given in fig. 8.

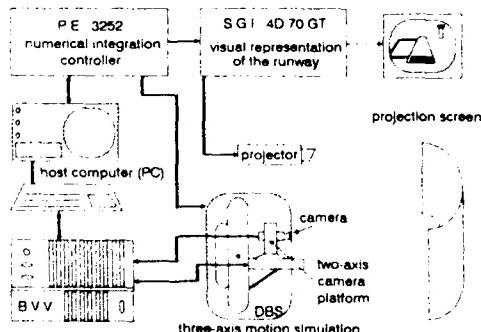


Fig. 7: Harware-in-the-loop-simulation

The advantage of HIL-simulation is, that on the one hand all sensors and data evaluation electronics (with their dirt effects, hard to be modelled), can be included in the investigation, but that on the other hand so called 'ground truth' data for the evaluation of system performance are readily available, since they are part of the numerical simulation. This type of simulator is widespread for both automatic guided weapon simulation and for pilot training; in dynamic vision, especially in the field of artificial intelligence, it is almost never used up to now. However, in several applications investigated at our institute it has

proven to be a valuable tool for efficient software and system development in this area also. With the interfaces between modules designed in the same way as they are in the flight hardware, the preparations for real flight tests on the remote airfield in Braunschweig could be kept to a minimum.

Especially in the area of testing feedback control laws with respect to wind and gust responses, the well defined and easily repeatable disturbances in simulation have definite advantages over the irregular, nonrecurring ones in the real world.

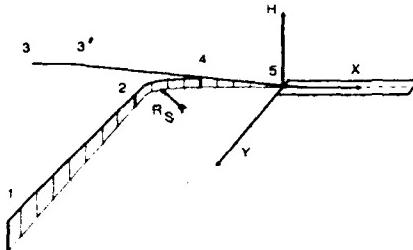


Fig. 8: Nominal trajectory shape

7. Test results

Closed loop performance with active control by the automatic visual guidance system could only be done in the simulation loop, since the airplane available from the Technical University Braunschweig did not have computer controllable actuators.

7.1 Closed-loop HIL-simulation results

The original flight hardware included in the simulation was: The CCD-TV-camera mounted on a two-axis (pan and tilt) platform for viewing direction control with angular rate feedback for inertial stabilization, the image sequence processing system BVV_2 as discussed in section 5, and the host-PC (80386) for system initialisation and data logging. Aircraft angular motion was realized by the DBS motion simulator for testing the viewing direction control platform; the inertial sensors in the aircraft have been simulated on the computer, however.

Wind effects have been incorporated by adding a (lateral) cross-wind component of -1 m/s starting at 800 m in front of the runway threshold; at 650 m an exponentially decreasing gust with a vertical component of 1 m/s and a cross-wind component of 2 m/s has been superimposed on the wind; the time constant of the gust was 6 s.

The simulation started at 950 m in front of the landing strip with a velocity of 50 m/s at an altitude of about 51 m; it ended at about 550 m down the runway with touch down of the landing gear. The ensuing transition to taxiing has not been investigated, since a completely different dynamical model would have been required both for state estimation and for control. It should be noted, however, that the vision sensor is well suited for

vehicle guidance both in this and the following ground roll guidance task along the taxways. The capability of performing these tasks has been demonstrated in essence with our road vehicle guidance program [2, 3, 5].

At about 250 m in front of the runway threshold the engines are throttled and the aircraft starts slowing down towards 43 m/s at touch-down; at the same time, the elevator is pulled in order to initiate the vertical flare decelerating vertical speed exponentially towards zero.

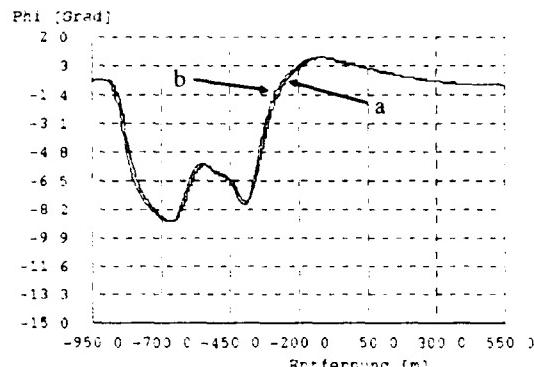


Fig. 9a: a) estimated b) simulated

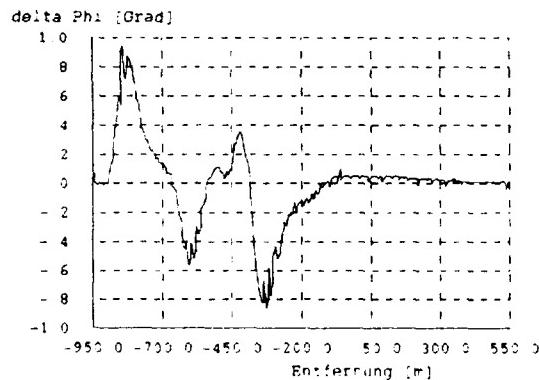


Fig. 9b: Difference between estimated and simulated

Fig. 9: Roll angle in degree

As long as the runway threshold is visible, the distance to it is estimated within 5 m accuracy; after losing the threshold out of sight at about 120 m distance from it, range x is merely predicted according to the model. The end of the runway can not be measured accurately enough from this low altitude to be of any use. Altitude above the runway could be estimated to within 1 m during approach and to less than half a meter over the runway; a rather large error briefly occurred during flare onset. This was due to the delay in the tilt viewing direction control loop at the sudden pitch rate change; it rather soon disappeared again.

The controller designed consists of a full state vector feedback with additional integral components for speed

and altitude, for which a strict time history profile is essential during landing approach. The effect of the vertical gust component is hardly noticeable in the trajectory flown.

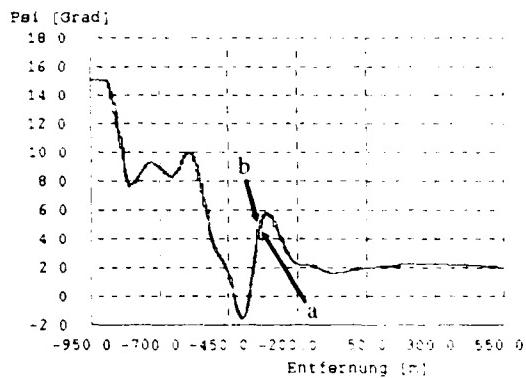


Fig. 10a: a) estimated b) simulated

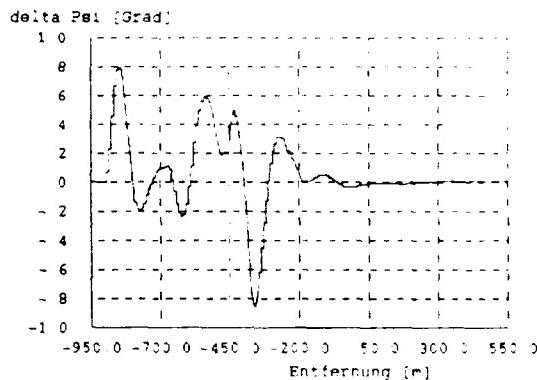


Fig. 10b: Difference between estimated and simulated

Fig. 10: Yaw angle in degree

In the lateral degrees of freedom the maneuver started with a lateral offset from the runway centerline of 85 m at 950 m distance and a heading error of 15 degrees towards the centerline. Within half a kilometer the lateral offset was reduced to less than 5 m, and within 0.75 km the heading angle settled at about 2 degrees relative to the runway, apparently mainly for crosswind compensation. During the initial maneuver, roll angles of up to 8 degrees occurred; the estimation errors in both roll and yaw angles were always less than one degree in magnitude (fig. 9, 10). The lateral gust induced an estimation error in lateral position of about 3 m for a short time; otherwise it was always less than one meter soon after initialisation (fig. 11). By adding another Kalman filter for runway width estimation the system was improved to be able to deal with not exactly known runway parameters; a 29 m wide runway could be validated to within 1 m accuracy. When the runway threshold was lost out of sight due to the close approach, the average estimated value was within 0.4 m of the true one.

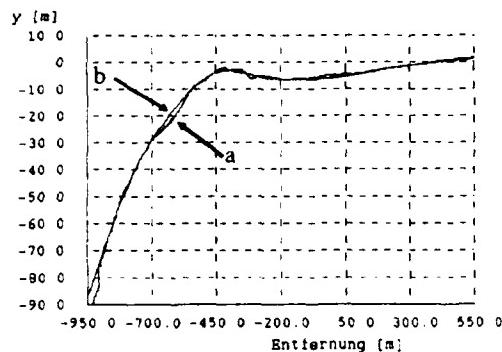


Fig. 11a: a) estimated b) simulated

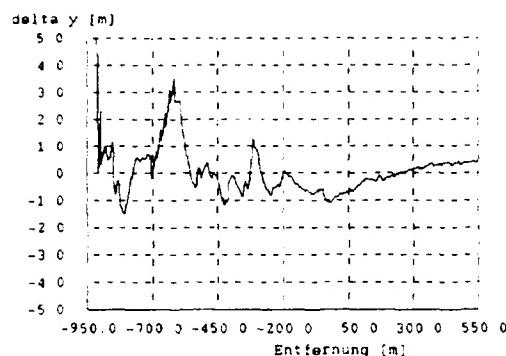


Fig. 11b: Difference between estimated and simulated

Fig. 11: Lateral offset y to the middle of the runway in meter

7.2 Flight test results in relative state estimation

After having gathered fundamental experience in simulation, test flights have been performed with the twin-turboprop airplane Do 128 of the University of Braunschweig. For evaluation of the accuracy of the optical

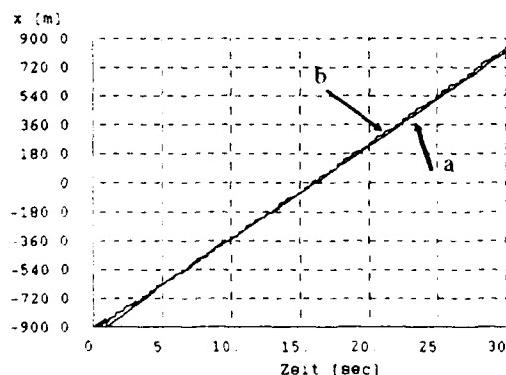


Fig. 12a: a) optical b) GPS system

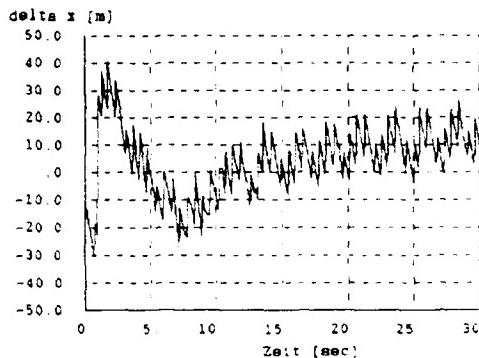


Fig. 12b: Difference between optical and GPS system

Fig. 12: Distance to the runway x in meter

results, a differential GPS system recorded the flight data in parallel to the optical system. The results in the translational degrees of freedom are given in fig. 12 to 14.

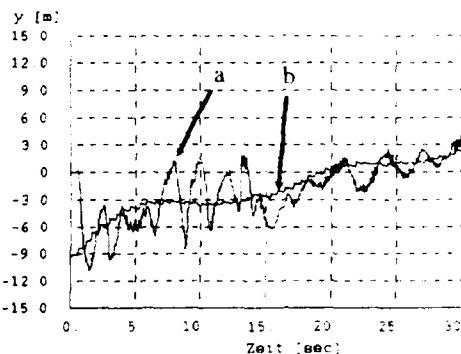


Fig. 13a: a) optical b) GPS system

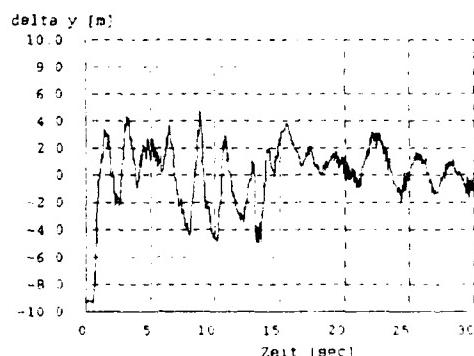


Fig. 13b: Difference between optical and GPS system

Fig. 13: Lateral offset y to the middle of the runway in meter

It can be seen clearly from fig. 13, that the optical evaluation process is getting better the nearer the airplane approaches the runway. The reason is, that the image of the runway is getting larger and the optical measurement values become better. For the evaluation of the altitude, the results of the GPS system showed errors. For comparison, a radiometric altitude-above-ground measurement system was used in addition, in order to record this value too.

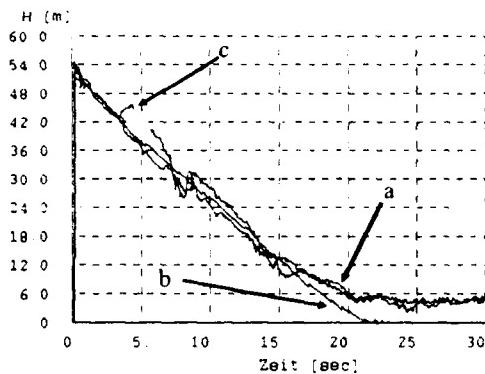


Fig. 14: Altitude in meter a) optical b) GPS c) radiometric

In order to even allow approaches to airports, for which the width of the runway is not known, an additive Kalman Filter was used to estimate this value. The width of the runway in Braunschweig is 29 m (difference between the white lane markings). Fig. 15 shows the result of the estimation, where the mean value estimated till about 200 m distance and 20 m elevation relative to the runway threshold is 29.1 m. After having lost this line within the image, the estimation is stopped and the mean value is taken.

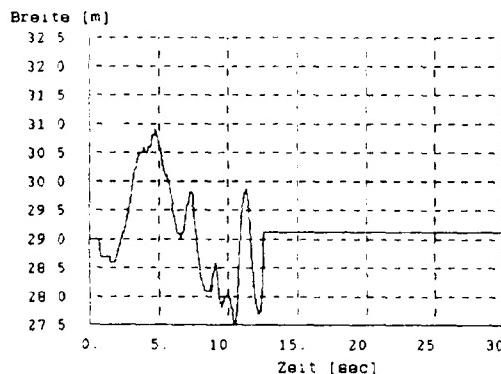


Fig. 15: Estimated runway width

8. Conclusion

Taking advantage of both spatial and temporal models for motion processes of objects, today's microprocessors, already, allow real-time image sequence processing, dynamic scene understanding and visual motion control in visually not too complex scenes in the 10 to 30 Hz range. The powerful microprocessors of the near future (200 MIPS-class), interconnected by high data rate communication networks, will enable the sense of vision for machines in well structured environments.

The method developed for video image sequence evaluation in the optical range may easily be adapted to other sensors like infrared, low-light-level TV or even imaging radar. With these, night vision or all-weather visual capabilities may become possible.

Combining visual and inertial sensor data evaluation has complementary beneficial effects: motion blur at high angular rates will deteriorate image processing; high rates, however, can easily be measured inertially at low cost. Inertial sensors become expensive when they have to be trimmed to long term stability. Static references for long term stabilization, however, can be measured easily by image processing, once the sensors have been roughly stabilized inertially. Therefore, good overall performance at low cost may be expected by a proper combination of both sensors.

A similarly beneficial effect may be achieved for navigation by combining vision with GPS: The latter provides a rough estimate of geographic coordinates, so that an intelligent vehicle capable of visual landmark navigation can start its visual search for landmarks in a rather small search area; accurate position determination relative to the landmark may then be achieved by visual tracking over time taking models for the ego-motion into account.

This new technology will allow autonomous on-board navigation and control capabilities unachievable up to now.

Fully autonomous control in a landing approach till touch-down has been demonstrated in a hardware-in-the-loop simulation in real time (16 Hz) including wind and moderate gusts. The same hard- and software has been installed in a twin-turboprop aircraft; within five days of testing, first results in relative state estimation during manually flown landing approaches have been obtained. These results compare well to differential GPS and radio altimeter results for the same flights.

References

- [1] Dickmanns E.D.: Dynamic Vision for Locomotion Control - an Evolutionary Path to Intelligence. CCG-Lehrgang SE 3.02, 1991.
- [2] Dickmanns E.D., Mysliwetz B., Christians T.: "Spatio-temporal guidance of autonomous vehicles by computer vision". IEEE-Trans.on Systems, Man and Cybernetics, Vol.20, No.6, Nov/Dec 1990, Special Issue on Unmanned Vehicles and Intelligent Robotic Systems, pp 1273-1284

- [3] Mysliwetz B.: Parallelrechner-basierte Bildfolgen-Interpretation zur autonomen Fahrzeugssteuerung. Dissertation an der Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, München, 1990.
- [4] Wünsche H.J.: Erfassung und Steuerung von Bewegungen durch Rechnersehen. Dissertation an der Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, München, 1987.
- [5] Dickmanns E.D., Graefe V.: a) "Dynamic monocular machine vision", b) "Application of dynamic monocular machine vision". J. Machine Vision Application, Springer-Int., Nov. 1988, pp 223-261.
- [6] Proskawetz K.O.: Ein Beitrag zur Genauigkeitssteigerung bei der Parameteridentifizierung nichtlinearer Prozesse am Beispiel der Flugzeugbewegung. Dissertation TU Braunschweig 1989.
- [7] Graefe V.: Dynamic Vision Systems for Autonomous Mobile Robots. IEEE Workshop on Intelligent Robots and Systems - IROS'89; Tsukuba, Sept. 89.
- [8] Maybeck P.S.: Stochastic Models, Estimation and Control, Vol.1, 2; Academic Press, New York 1982
- [9] Schell F.-R.: Bordautonomer automatischer Landeanflug aufgrund bildhafter und inertialer Meßdatenauswertung. Dissertation an der Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, München, 1992.
- [10] Dickmanns E.D., Zapp A., Otto K.D.: "Ein Simulationskreis zur Entwicklung einer automatischen Fahrzeugführung mit bildhaften und inertialen Signalen". In Breitenecker, e.a. (Hrsg.); Simulationstechnik; Informatik-Fachberichte 85, Springer-Verlag, 1985
- [11] Otto K.D.: Linear-quadratischer Entwurf mit Strukturvorgaben. Dissertation Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, München, 1990
- [12] Bierman G.J.: Factorization Methods for Discrete Sequential Estimation. Academic Press, New York 1977.

**Scene Correlation for INS Aiding in Flight-Test Systems
- Runway-Referenced Flight-Tests with On-Board Sensors only**

B. Stieler
H.-U. Dohler
German Aerospace Research Establishment (DLR)
Institute of Flight Guidance
3300 Braunschweig, P.O. Box 32 67, FRG

Abstract

Inertial and image-derived measurements for runway-referenced flight path computations are investigated. They open the way for flight-tests without ground-based sensors and with a minimum of a priori knowledge or none at all about the runway position and direction in absolute coordinates. They are ideally suited for inspecting landing aids at congested civil airports, for instance.

The investigations concentrate on the problems of the baroinertial and inertial altitude measurements, on the achievable accuracies of image-derived measurements and their integration with inertial measurements. It is shown that glidepath and flight track accuracies in the order of 0.01 deg (1σ) should be achievable.

1. Introduction

In many cases flight-tests are carried out near an airport where trajectory measurements with respect to the runway (RWY) are of interest. Flight inspection systems for testing landing aids of civil airports may serve as an example.

Traditionally these flight-tests are based on sensors on board and on ground as indicated in Fig. 1.1. A microwave radar or a laser tracker is installed on ground with a known absolute position reference and measures the relative position of the aircraft with respect to its absolute position. This is transmitted to the aircraft for aiding the inertial navigation system (INS) navigating in absolute coordinates, too. As next step the flight path with respect to the absolute position of the runway reference point and direction is computed.

With fixed installations on ground this procedure is in principle a very good solution. The "Avionic Flight Evaluation System (AFES)" of DLR works in this fashion [Stieler 91/1]. The burden of positioning and orienting the ground-based sensors is carried out only once. Errors in their absolute position and direction are easily uncovered. They may be cumbersome when data of flight inspections on a plurality of airports are evaluated with submeter accuracy requirements.

The advent of the "Global Positioning System (GPS)" alleviates the burden just mentioned. Orienting the reference station on ground as basis for differential GPS measurements is no longer required, but the data evaluation is in principle the same as indicated above.

Human beings do not orient themselves within the surrounding in absolute coordinates, but with the help of their eyes in relative positions. Methods of digital imaging offer similar possibilities for orientation and data evaluation. Their combination with an INS and memory capacity for data storage allows the flight path estimation back into the past or forward into the future, which works in a similar way as the human being does with his optical sensors and memory.

This paper deals with runway-referenced flight-tests based on on-board sensors only -especially optical and inertial sensors.

Experiences with a flight inspection system of the German Aviation Authority "Bundesanstalt für Flugsicherung (BFS)" are outlined first [Stieler 91/2]. It is based on multi DME, and INS for testing en route radio aids as VOR/DME stations. For testing landing aids it is based on a downward looking CCD camera, INS plus radio and baro altimeters.

Possible improvements and their effect on the achievable accuracy of the flight path reconstruction are discussed. The paper is concluded with an outlook for runway-referenced flight-tests without a priori knowledge about the runway position and direction in absolute coordinates.

2. Experiences with the BFS Flight Inspection System

The "Gemeinsame Flugvermessungsstelle (GFMS)", a branch of BFS, is responsible for testing en route air traffic navigation systems and landing aids in Germany.

The "Digitales Meßsystem (DMS 88)" for testing enroute air traffic systems consists of an INS aided by multi DME. For testing landing aids DMS 88 is in addition equipped with a downward looking CCD-camera and a radio altimeter for taking measurements during the runway overflight. The reference flight path, requiring an uncertainty for ILS localizer and glidepath measurements below 0.02 deg., is reconstructed as indicated in Fig. 2.1. The data storage of the INS, aided in the horizontal channels by multi DME and in the vertical channel by a barometric altimeter, is initiated prior to passing the outer marker. When the aircraft flies over the patterns at the beginning and the end of the runway, the CCD camera takes one position fix each time. In addition the altitude is measured by means of the radio altimeter. The measurements serve to aid the absolute horizontal and vertical INS positions using a forward Kalman-filter. After passing the second pattern at the end of the runway, the flight path for the approach phase is reconstructed by backward filtering the stored flight path including all external measurements.

DMS 88 thus opens the way for inspecting landing aids at congested civil airports without ground installations and major interference with the traffic flow. During the test campaigns two items for possible improvements within the system were uncovered. They concerned the glidepath computation and the reliability in obtaining the required two position fixes from the downward looking camera.

The following sections deal with flight-test, laboratory test and simulation results backing up steps for such improvements. Changes in the data evaluation concept would in addition allow to carry out landing aid inspections at airports with unknown position and direction of the runway.

Due to their independence from ground installations and their error characteristics the combination of INS plus forward-looking camera is ideally suited for this purpose. The goal is to obtain an accuracy of 0.01 deg for the glidepath and flight track computations. The corresponding handling of errors in the submeter level requires a close look into the performance of the INS and the camera, which will be treated in the next two chapters.

Before going into details a word has to be said about the integration of satellite navigation (GPS) receivers into a flight inspection system. GPS position measurements in the submeter level require a reference station on ground. The goal of this paper is to show that the mentioned glidepath and flight track accuracies can be achieved without ground installations. It is out of question that a future flight inspection system will make use of GPS, primarily as replacement or supplement of multi DME for enroute flight path computations. If differential GPS measurements are available from reference stations permanently installed on ground the situation is different. They will then provide a very useful supplementary source of information for the runway-referenced flight-tests investigated in this paper.

3. Critical Review of the INS Vertical Channel in Comparison to the Horizontal Channels

Because of its inherent instability [Winter 75, Stieler 82] the INS vertical channel is barometrically aided as indicated in Fig. 3.1. Roughly speaking it has the following characteristics.

Due to the fairly weak barometric coupling the

- accelerometer measurements dominate the short-time accuracy and
- the barometric altitude and its bias dominate the long-time accuracy.

But what is short-time and long-time for an aircraft approaching the runway? Which is the barometer bias in this phase?

The biggest problem for an accurate barometric altitude result is the measurement of the static pressure in the moving aircraft. Depending on the position of the static pressure holes in the fuselage, this measurement depends more or less on air flow parameters such as [Wuest 80]

- the magnitude of the air flow, i.e. the true airspeed (TAS)
- the direction of the air flow, i.e. the angles of attack and sideslip
- the flap and throttle settings

Test aircraft are mostly equipped with a noseboom to minimize these effects.

The conversion of the pressure into an electric signal is affected by the length of the pressure tube, which defines the time lag in the order of magnitude of 1 s [Redeker 85, Nord Micro 84]. For the computation of barometric altitude above the runway from static pressure measurements additional parameters have to be known

- ground pressure and temperature and
- outside temperature

Differences in these parameters cause additional bias and scalefactor variations

If the barometric altitude error remained constant during the approach phase an accurate glidepath computation would be possible based on a few additional external measurements. Unfortunately this is not the case. Fig. 3.2 shows flight-test results obtained with the twin engine turboprop aircraft DO 228 of DLR with static pressure holes in the fuselage. The differences between the barometric altitude obtained from the ADC and the AFES laser tracker are shown here. They show a very problematic time characteristic during the approach phase which will not be removed by feeding the barometric altitude into the inertial system for obtaining the baroinertial altitude.

In the GFMS test aircraft which has its static pressure holes in the fuselage, too additional effects due to changing sideslip angles were observed.

Using a noseboom for static pressure measurements would reduce these effects considerably, but it seems hopeless to aim at a submeter accuracy in the baroinertial altitude computation.

The situation is quite different for the pure inertial altitude. Though it is inherently unstable as already mentioned, its error growth is modelable over a certain period of time as shown below and it can serve as a basis for an aided altitude and glidepath computation. The problem in this approach is to get access to it in a commercially available INS. Since the customer cannot change the INS software, the feedback gains in Fig. 3.1 cannot be put zero without a costly manipulation by the manufacturer. The second possibility of double integrating the vertical acceleration as basis for inertial altitude computation is burdened with the fact that it is only accessible in the commercial INS after having passed a first order Butterworth filter, implemented with the following recursive formula in the Litton LTN-90 [Litton 83]

$$\begin{aligned}
 (3.1) \quad a^*(n) &= c(n-1) + 0.5 [c(n) - c(n-1)] \\
 c(n) &= c(n-1) + k [a(n) - c(n-1)] \\
 k &= 2Dt/(Dt + 2T) \\
 a^* &= \text{filtered vertical acceleration} \\
 a &= \text{measured vertical acceleration} \\
 T &= 0.5 \text{ s}
 \end{aligned}$$

In our laboratories we investigated 2 methods of extracting the inertial altitude from this system.

Method 1

Using the INS altitude and feeding it back as barometric altitude into the system. From Fig. 3.1 it can be easily derived that the feedback signal is zero in this case and the INS altitude should be the pure inertial altitude.

Method 2

Deriving the inertial altitude from the vertical acceleration after passing it through an inverse filter. The straightforward procedure for this method is to solve Eq. (3.1) for the acceleration $a(n)$ measured. This procedure requires that the INS output is sampled synchronously with the INS computer clock frequency of 64 Hz. For the laboratory setup with 20 Hz sampling frequency the following equation gave the best results

(3.2) \quad a(n) = a^*(n) + [a^*(n) - a^*(n-1)] T \cdot \Delta t

During the laboratory tests the LTN-90 was exposed to linear motions of ± 20 cm amplitude at 1 Hz frequency approximately. As to be seen on Fig. 3.3 the rocking table was inclined by 25 deg with respect to the vertical and 45 deg with respect to true north. So all 3 INS channels sensed a component of the linear oscillation.

Fig 3.4 shows over 30 s of time for both methods the residues between the reference altitude ($h = 0$) and the inertial altitude after subtracting a 2nd order polynomial for compensating the typical inertial errors. In Fig. 3.5 the inertial altitude for both methods is plotted over 4 min. The plots and the polynomial coefficients prove that both methods give results with excellent correspondence. With a 3rd order polynomial the residues become much smaller, as shown in Fig. 3.6 for two different tests.

For a longer period of time, e.g. 15 min, the errors of the inertial altitude, computed as outlined above, cannot be modelled as well as can be seen from Fig. 3.7a. The peaks of the residues with respect to a 3rd order polynomial are ± 1.5 m, approximately. In a simulation it was found out that a random walk coefficient of $r = 0.2$ (m/s)/ v_h generates residues of similar magnitude (see Fig. 3.7b). This random walk corresponds to an acceleration signal with 0.005 g resolution at 20 Hz sampling rate. Since the LTN 90 specifications quote $1.22 \cdot 10^{-4}$ g at 64 Hz, the reason for this problem lies in the asynchronous data collection of the laboratory setup. There is no physical reason why within the INS the vertical channel should be inferior to the horizontal channels [Knickmeyer 90]. Fig. 3.8 shows the corresponding laboratory test results for the horizontal residues computed from the north and east velocities. The nonlinearity at 60 s is due to the limited resolution of 0.125 kn of the LTN 90 output signals. If the system is standing still, this may generate a position error slope of equal magnitude which vanishes when the system is moving. The random walk coefficient at 16 Hz update rate is $r = 0.27$ m/ v_h , which is negligible. The higher noise in the horizontal position with respect to the inertial altitude is due to the difference in the LTN 90 transport delay which is 60 ms for the accelerations and 110 ms for the horizontal velocities.

Both methods outlined above for inertial altitude computation give good results for mission times in the order of 4 min, which is sufficient for the approach phase. Method 2 is the preferred way for computing the inertial altitude, because

- the baroinertial altitude computation is not interrupted and
- due to the inertial altitude error growth Method 1 works for 10 min only, approximately, when the computed vertical velocity does not exceed limits so that the internal INS built-in test equipment (BITE) classifies it as "invalid".

In Method 2 we have found a basis for inertial altitude measurement. For on-line and off-line glidepath computations external measurements for aiding the inertial altitude are necessary. There are only 2 on-board sensors for the measurement of the altitude with respect to the runway.

- the radio altimeter which measures the relative altitude on-line, i.e. requires the storage of the runway profile as a priori knowledge, and
- the forward looking CCD camera which measures the relative altitude already during the approach phase, i.e. it requires the relative altitude difference between both runway patterns as a priori knowledge only.

We will concentrate in the following on the second sensor

DLR's concept for scene correlation navigation is discussed in the next section which is followed by accuracy considerations for the position computation in the approach phase based on the runway pattern.

4. Concept for the Scene Correlation Navigation

In Fig. 4.1 a block diagram for the scene correlation navigation is shown. It consists of three main parts: image processing, model generation and identification with position and attitude estimation.

With the help of information about position and orientation, delivered by the inertial navigation system (INS), moreover under consideration of foreknowledge about errors of these data, objects which can be detected by the scanning device are selected from a digital terrain model (DTM). These objects are projected from 3-D space into the 2-D image plane using a hierarchical data structure (scene generation, artificial scene description).

Image processing is done in two steps. First, preprocessing clears the image from brightness fluctuations and nonlinear geometrical distortions depending on the sensor. In a second step (segmentation) features are extracted and stored in the same way as the transformed DTM objects. Depending on the expected image procedures of pattern recognition are chosen from a stock of algorithms (image processing knowledge method selection). In consideration of the applied segmentation procedures and with the help of a selection step (scene evaluation) a subset of objects contained in the scene description is produced which offers a reliable image interpretation (image model).

After assigning the features that have been extracted from the aerial image to the image model (object identification) the real position and attitude angles of the aircraft are calculated (photogrammetry) and can be used as update information for the INS-data.

4.1 Model Generation and Photogrammetry

A digital map on board (digital terrain model, DTM) contains data about the structure of terrain objects. Each object is described by its geometry and attributes, e.g. the type of the object, its class and name. The object's geometry is expressed by the position of the vertices of surrounding polygons in world coordinates (Gauss-Krueger coordinates) and the height above MSL, using $X_w = (x_w \ y_w \ z_w)^T$. For experimental purpose a DTM from the region around the Braunschweig airport was digitized (size $6 \times 4 \text{ km}^2$), based on the Deutsche Grundkarte (Scale 1:5000). In the near future geographical information systems (GIS) for example the Amtliches Topographisches Kartographisches Informationssystem (ATKIS) of the German Federal Surveying Offices (Landesvermessungsamt) are available.

In order to derive the image model by using the DTM data, the objects are transformed to the image coordinate system $X_i = (x_i \ y_i \ z_i)^T$. Furthermore only such DTM objects are chosen which are sufficiently recognizable with respect to the available image processing methods. This transformation is controlled by the attitude and heading angles of the aircraft and its approximate position $X_a = (x_a \ y_a \ z_a)^T$. The camera is described by a simple pinhole model (nonlinear geometrical distortions are already eliminated) with focus length c and the mounting angles (roll, pitch and yaw) inside

the aircraft. The real object positions in the image coordinate system are determined by the hypothesis driven assignment of the objects to the features contained in the processed image.

Using well-known procedures of photogrammetry [Konecny 84], the parameters of the real transformation are determined from the corresponding points of the objects in the image and world coordinate systems. These parameters are elements of a redundant nonlinear set of equations. They can be linearized and solved iteratively by the least square method. Thus the actual position of the aircraft $X_a = (x_a, y_a, z_a)^T$ is available as update information for the INS. The same method is suitable for determination of the aircraft's attitude angles and for camera calibration i.e. the determination of the mounting angles and the focal length.

4.2 Image Processing

In a preprocessing step sensor distortions are corrected by different procedures of filtering and transformation. Besides global grey scale operations such as histogram equalization also local operators are taken into consideration. Regarding a sensor specific model these operators can compensate a position dependent sensitivity or defective cells of CCD sensors. Moreover geometrical distortions are equalized if necessary.

In a second step (segmentation) the image is decomposed into significant parts and described by features of different kinds. The selection of the applied segmentation methods and their parameters from a stock of image processing algorithms (IP expert knowledge) is controlled by the expected image contents (IP expert system).

Besides standard local procedures for the extraction of linear features [Dunham 86], also alternative approaches are made. Therefore methods from the field of graph search techniques like dynamic programming [Besslich 89 Martelli 76] and heuristic contour following techniques [Montanari 71] can be used. These procedures are very profitable because they allow the inclusion of knowledge concerning the model of special image structures using grey scale images or derived gradients as well. Moreover not only linear but also region like features are extracted by segmentation using additional methods of texture analysis.

Features and segments extracted by these procedures are stored for further processing in a hierarchical data structure. In addition method parameters and quality criterions are attached to each feature as attributes. The use of IP expert knowledge makes this technique fairly flexible.

4.3 Object Identification

The object identification is done by the assignment of extracted features to the objects of the image model. In the literature many procedures for solving this problem are published, for example methods used in the fields of workpiece recognition, robot vision etc [Salzbrunn 90 Tropf 80]. The selected method allows a fault tolerant, displacement and rotation invariant 2-D feature assignment. In addition a slight scale adjustment is made. This procedure, a modified A-algorithm [Nilsson 82], builds up valued concurrent hypotheses for the objects contained in the image model. This corresponds to the procedures of the so-called analysis by synthesis described in [Tropf 80]. For the generation of a starting hypothesis a feature of the image model is used estimating the parameters of rotation and displacement for the assignment (for example a corner, i.e. the point of

intersection of neighbouring straight line segments). Starting with these initial hypotheses more features of the image model are searched. The validation of the hypotheses depends on the similarity of the compared features regarding a given threshold of tolerance for position and orientation. The parameters of displacement and rotation are calculated for each new feature so that the global error of position of all object features is minimized. After assigning all model features to some image features the verification process will be terminated. If the global matching quality is higher than a recognition threshold the model is assumed to be recognized. In this case all matched image features are marked to be used. All other hypotheses built upon these features are devaluated. If a hypothesis quality sinks below the recognition threshold it will be rejected.

5. Accuracy Considerations for the Runway Approach as an Example

The position accuracy of image based navigation systems severely depends on the actual geometric arrangement of the world objects and the viewing camera. Therefore it is impossible to give an overall estimate of the resulting position error. In the case of many objects which are spread over the whole image plane it is obvious that the achieved accuracy will be much greater than in the case of only one single object of small size somewhere in the image. It should be emphasized that in general the position error depends on the image scale, i.e. the nearer the objects appear and the longer the focal length of the camera, the more accurate the determined position will be. This is the main difference between image based navigation and other navigation methods (GPS or DGPS) which deliver a position independent accuracy. Especially for the regarded case of only one single object in the image field, i.e. the stripe of the runway, it is necessary to make some detailed assumptions for estimating the resulting errors.

The estimation of the resulting position error during the landing approach was done via different computer simulations. In these simulations different kinds of error sources were assumed which can be regarded as system constraints. These error sources are:

- 1) errors in the INS angles
- 2) a focal length error of the camera or noncalibrated nonlinear distortions of the camera lens and
- 3) random errors of image coordinate locations of reference points resulting from the approximation of straight line segments to the contours in the image plane

For these simulations a fixed camera model is used. The forward-looking camera is mounted at a fixed position and orientation with respect to the aircraft axes so that for the aircraft approach the baseline of the runway is located in the center of the image plane. The simulated camera target has a size of 36 x 24 mm and is digitized with 360 x 240 pixels (pixel size 0.1 mm). The applied focal length is 50 mm. These data can easily be transformed to commercially available CCD cameras. For example, a focal length of about 12 mm for the KODAK-MEGAPLUS with a pixel size of 6.8 x 6.8 microns and a resolution of 1280 by 1024 pixels would approximately achieve the same viewing angles (however in this case the resolution is about 4 times higher). The simulation is done at different points during the landing path, i.e. at 1000, 500, 250, 100 and 50 m before of the intersection point of the flight path and the runway. For all these points the exact locations of reference points in the image plane are computed by projection (forward com-

putation). After adding the regarded errors to the recording parameters and the image locations of reference points the position computation is carried out (backward computation). An example for the estimated positions is shown in Fig. 5.1. The aircraft's position has a distance of 50 m with respect to the baseline of the runway. The calibration error is assumed to be 0.1%, and the random location error in the image plane is assumed to have a standard deviation of 0.5 pixel (so that in 95.5% of all cases the location differs not more than 1 pixel from the true location). For this scatter diagram 1000 random data are taken into account. The resulting position error (mean and standard deviation) can be seen in Fig. 5.1. These scatter diagrams are computed for all five regarded distances from the runway under two different assumptions.

First it is assumed that the aspect angles are taken from the measured INS angles (with a systematic error 0.1 degrees) and that only the corner points between the baseline and the edges of the runway are extracted from the image (see Fig. 5.2 b)). The results are listed in Table 5.1.

Second it is assumed that the aircraft position and its attitude angles are computed from three image points horizon point from the intersection of the runway edges and the left and right corner points of the baseline (see 5.2 c)).

In the next section the results listed in Table 5.1 will be applied to the computation of a runway-referenced flight path.

6. Results for a Runway-Referenced Flight Path Computation Based on INS Data and Image Position Fixes

For the runway-referenced flight path computation the absolute positions of the runway patterns need not be known. The first pattern is used as reference position. The known relative position of the second pattern with respect to the first one lays down the horizontal and vertical reference directions as shown in Fig. 6.1.

Fig. 6.2 shows the block diagram for the runway-referenced flight path computation. The time 0 is a point of time prior to passing the outer marker. The integration of the horizontal velocities and the vertical acceleration are initiated. For the latter the baroinertial altitude serves as initial condition. For the following discussion it is assumed that this altitude has an initial bias of $\delta h_b = 100$ m with an uncertainty of $\sigma(\delta h_b) = 5$ m for an aircraft equipped with a noseboom and $\sigma(\delta h_b) = 30$ m in other cases. These uncertainties take into account the bias difference between the initiation and the runway overflight. The initial horizontal position can be assumed zero with an uncertainty of say $\sigma(ds) = 100$ km. The horizontal velocities are integrated in the runway-referenced coordinates shown in Fig. 6.1.

During the approach the camera measurements are used for aiding the inertial position in a Kalman filter. They are initiated at a distance of 1000 m ($f = 50$ mm) or 2000 m ($f = 100$ mm) before the runway patterns (see Table 5.1). After having passed the second pattern, the backward filter is initiated as indicated in Fig. 6.2.

For the rest of the paper we concentrate primarily on the vertical channel only, which, as shown above, is more difficult to handle than the horizontal channels. The following results are based on true INS data collected on the rocking table shown in Fig. 3.3 and on simulated position fixes. The σ_2 values of Table 5.1 served to distribute randomly these fixes with respect to the reference

flight path, $n = 0$ in our case. They were obtained for a focal length of $f = 50$ mm. In order to simulate the focal length $f = 100$ mm, the corresponding distances in Table 5.1 were multiplied by 2. Random synchronization errors between INS and camera were partially taken into account by limiting the σ values to 0.1 m.

For an approach of 4 min duration the vertical channel was modelled in the Kalman filter with the following state vector:

$$(6.1) \quad \mathbf{x} = (\delta h \ \delta v \ \delta a)^T$$

with

δh = inertial altitude error, δv = inertial velocity error, δa = acceleration error

The initial covariances and random walk coefficients were chosen as

$$(6.2) \quad \begin{aligned} \sigma(\delta h) &= 5 \text{ m or } 30 \text{ m} \\ \sigma(\delta v) &= 10 \text{ m/s} \\ \sigma(\delta a) &= 0.01 \text{ g} \\ r_i(\delta h) &= 0.001 \text{ m/s}^2 \\ r_i(\delta v) &= 0.2 \text{ m/s} \cdot \text{s} \\ r_i(\delta a) &= 0.0001 \text{ g} \cdot \text{s}^2 \end{aligned}$$

The investigations were carried out in steps with decreasing a prior knowledge:

- the relative positions of the runway patterns i.e. the runway length and the altitude difference are known and
- the relative positions of the runway patterns are unknown

6.1 Test Results Using Measurements with Respect to Runway Patterns of Known Relative Positions

For known position differences between the two runway patterns i.e. known distance and relative elevation, the image measurements serve as position measurements in the coordinate system shown in Fig. 6.1.

The test results are presented in Figs. 6.3 to 6.5. They show the altitude estimation residues for different runs and different INS data files. They are zero approximately when the aircraft is close to the patterns at the beginning and the end of the runway and they increase backward in time due to the characteristics of the camera and due to the fact that the initial baroinertial altitude measurement is the only external measurement from $t=0$ up to the time when the runway pattern is in sight of the camera. The dotted curves are the $\pm 1\sigma$ bands computed by the backward filter. The shaded cones mark over a distance of 10 km the ± 0.01 deg glidepath limits at an approach speed of $V = 70$ m/s.

Figs. 6.3 and 6.4 show the effect of the camera focal length and the runway length i.e. the distance between both runway patterns on the altitude estimation accuracy. It is assumed that the initial altitude is equal to the baroinertial altitude at $t = 0$ and that the baroinertial bias is estimated based on the camera measurements. The uncertainty between the inertial altitude initiation at $t = 0$ and the baroinertial bias estimation is $\sigma(\delta h_b) = 5$ m (assuming a noseboom-equipped aircraft). From the results one may conclude that with a focal length of $f = 100$ mm the 0.01 deg 1 σ glidepath accuracy can be met.

If the baroinertial altitude uncertainty is bigger, i.e. $\sigma(\delta h_b) = 30 \text{ m}$, as it may be the case for a none nose-boom-equipped aircraft (s. Fig. 3.2) the glidepath accuracy is heavily affected as to be seen from Fig. 6.5. This gives an indication of the benefits of good barometric measurements. It should be kept in mind that they are not used for an accurate absolute altitude computation in the submeter range, but for providing the inertial altitude computation with a starting point whose bias should not change above the $\sigma(\delta h_b)$ level during the approach.

Fig. 6.6 shows the results for two successive runway approaches with a time difference of 7 min approximately. Here and in the following investigations the camera focal length is $f = 100 \text{ mm}$ and the runway length is 4.5 km. Due to the longer mission time of 1/4 h the state vector was augmented by the acceleration slope δb .

$$(6.3) \quad \dot{x} = (\delta h \ \delta v \ \delta a \ \delta b)^T$$

with the following initial covariance and random walk coefficient.

$$(6.4) \quad \begin{aligned} \sigma(\delta b) &= 0.0005 \text{ g/h} \\ r(\delta b) &= 0.0001 \text{ (g/h)}\sqrt{\text{h}} \end{aligned}$$

Both plots in Fig. 6.6 prove again the benefits of an accurate baroinertial measurement as initial condition. The results of the 2nd runway overflight do not depend on this measurement any longer.

6.2 Test Results Using Measurements with Respect to Runway Patterns of Unknown Relative Positions

If the aircraft carries out flight-tests with respect to a runway of unknown absolute and relative positions of both patterns it has to fly a traffic pattern, i.e. two approaches and runway overflights for achieving the required glidepath and flight track accuracies. The pattern size is the only *a priori* knowledge then required for measuring the aircraft position x, y, h with respect to the first pattern in the runway coordinate system (s. Fig. 6.1) and for measuring its y -component with respect to the second pattern. The latter is its offset from the runway direction which certainly can also be measured with respect to the runway center line or the runway edges.

If stereo-optical measurements are taken, even the runway pattern size is no longer a prerequisite for measuring the relative aircraft position with respect to the first pattern and for measuring its course offset. In this case no *a priori* knowledge about the runway is required for the runway-referenced flight path computation.

The results in Fig. 6.7 are surprisingly good. The glidepath accuracy of 0.01 deg (1 σ) can still be met, even for the case of an initial baroinertial altitude bias uncertainty of $\sigma(\delta h_b) = 30 \text{ m}$. During the first approach this is true for a short distance before the first runway pattern. For greater distances the glidepath accuracy exceeds the $\pm 0.01 \text{ deg}$ limits.

At the end of these investigations we will cast a brief glance at the horizontal channels. For the following results the system error model was left unchanged with respect to Eq. (6.3), i.e. the Schuler loop was neglected for the mission time of 1/4 h. Also the coupling between both horizontal channels due to the changing heading was not taken into account. The following parameters were set to new values:

$$(6.5) \quad \begin{aligned} \sigma(\delta s) &= 100 \text{ km} \\ r(\delta s) &= 0.2 \text{ m/s} \end{aligned}$$

$$r(\delta v) = 0.01(\text{m/s})\sqrt{\text{h}}$$

The fairly big value for the random walk on the position level is due to the north and east velocity resolution of 0.125 kn at a sampling frequency of 20 Hz within the LTN 90. For the image-derived position fixes the σ -values of Table 5.1 were used, with a limitation to 0.1 m again to take into account synchronization errors between INS and camera.

Fig. 3.8 shows the estimation residues in the x and y directions. Due to the characteristics of the image-derived position fixes the along track error δs_x cannot be estimated as well as the cross track error δs_y . This is no problem, since glidepath and flight track accuracies do not depend so much on distance. The cross track error defines the flight track accuracy. The shaded cones in Fig. 3.8 prove that the 0.01 deg (1 σ) accuracy level can be reached easily.

7. Summary

Runway-referenced flight-tests with on-board sensors only are optimally carried out based on inertia and image measurements of the patterns at the beginning and the end of the runway.

Experiences with the flight inspection system DMS 88 of the German Aviation Authority indicate that the glidepath computation based on baroinertial measurements is critical (Chapter 2). Barometric bias trends during the approach phase are hardly modelable.

For short periods of time the purely inertially derived altitude seems to be more appropriate, but it is difficult to get access to corresponding accurate measurements in a conventional INS. Two different methods and their laboratory test results are outlined in Chapter 3.

Chapter 4 deals with DLR's concept of image-derived navigation and Chapter 5 with the achievable accuracies of position of measurements based on runway patterns of known size. Not only position updates are possible but also attitude and heading updates using photogrammetric relations.

Based on simulated image-derived position updates and true inertial measurements the accuracy of the flight path computation - primarily the glidepath computation - is analyzed in Chapter 6. The data handling is similar to the one in DMS 88 during the approach and runway overflight. All data integrated in a forward Kalman filter are stored and evaluated in a backward filter after passing the runway end.

For known size and relative positions of both runway patterns it should be possible to reconstruct the flight path with 0.01 deg 1 σ -accuracy. A camera focal length of $f = 100 \text{ mm}$ is preferable. The results indicate also the importance of a fairly stable baroinertial altitude measurement as initial condition of the inertial glidepath measurement. Stable means that the barometer bias of unknown magnitude at the outer marker should not change more than 5 m with respect to the runway overflight. When the runway is overflown twice in time after the test aircraft has flown a traffic pattern the glidepath accuracy for the second approach does not depend on barometric altitude accuracy any longer.

The repeated runway overflight opens the possibility of an accurate runway-referenced flight path computation without the relative position of both patterns and only their size as *a priori* knowledge. Stereo-optic measurements on-board the aircraft allow to compute the run-

way-referenced flight path without any a priori knowledge

The first runway pattern overflown twice serves to calibrate the along track and the altitude, and the second pattern, the runway centerline or the runway edges serve to calibrate the runway-referenced direction. For glidepath and flighttrack computations a 10-accuracy of 0.01 deg should be achievable.

Flight tests with a gyro-stabilized, coarse resolution IR camera mounted in the nose of the DLR test aircraft DO 228 are evaluated at present (see Fig. 1.1). Further flight-tests will be based on the KODAK-MEGAPLUS Camera mentioned in Section 5.

Acknowledgement

The authors wish to express their gratitude to Mr. Zenz, Mr. Linde and Mr. Gruttemann for providing the flight and laboratory test data, to Mrs. Kraul and Mrs. Buhrig for typing the text and to Mrs. Stieler for revising the English text.

References

- [Besslich 89] Bassmann, H. Besslich, Ph W. Konturorientierte Verfahren in der digitalen Bildverarbeitung. Berlin, Heidelberg, New York Springer, 1989
- [Dunham 86] Dunham, J G. Optimum Uniform Piecewise Linear Approximation of Planar Curves. IEEE Trans Pattern Analysis and Machine Intelligence. Vol PAMI-8, 1986 pp 67-75
- [Knickmeyer 90] Knickmeyer, E. Accuracy of Gravity Vector Recovery with the LTN 90-100 RLG Strapdown System. Proceedings of the Symposium No. 107 Kinematic Systems in Geodesy, Surveying and Remote Sensing. Banff Alberta, Canada Sept 1990
- [Konecny 84] Konecny, G. Lehmann, G. Photogrammetrie. 4 Aufl. Berlin New York Walter de Gruyter 1984
- [Litton 83] Technical Description of LTN-90 Inertial Reference System. Litton Aero Products revised January 1983
- [Martelli 76] Martelli, A. An Application of Heuristic Search Methods to Edge and Contour Detection. Comm ACM 19, 1976, pp 73-83
- [Montanari 71] Montanari, U. On the Optimal Detection of Curves in Noisy Pictures. Comm ACM 14, 1971 pp 335-345
- [Nilsson 82] Nilsson, N J. Principles of Artificial Intelligence. Berlin Heidelberg, New York Springer, 1982
- [Nord Micro 84] Digital Air Data Computer. Part No 12238-3. Maintenance Manual with Illustrated Part List. 34-10-10 Oct 15/84
- [Redeker 85] Redeker, A. Vorsmann, P. Precise Vertical Speed Reconstruction Based on Vertical Acceleration and Barometric Altitude. Z. Flugwiss Weltraumforsch 9, 1985, Heft 4
- [Stieler 91 1] Stieler, B. Hurraß, K. Boden- und satellitengestützte Flugbahnmessung am Beispiel des Avionik-Flugerprobungssystems (AFES). Sammelband des Wissenschaftlichen Kolloquiums des Instituts für Flugfahrt der DLR Braunschweig, 04-05-06 1991
- [Stieler 91 2] Stieler, B. Hurraß, K. Bordautonome Flugbahnmessung am Beispiel eines ILS-Vermessungssystems. Sammelband des Wissenschaftlichen Seminars des Instituts für Flugfahrt der DLR Braunschweig 1991
- [Stieler 82] Stieler, B. Winter, H. Gyroscopic Instruments and their Application to Flight Testing AGARD-AG-160-Vol 15, 1982
- [Salzbrunn 90] Salzbrunn, R. Behnke, K. Konturbasierte fehlertolerante Erkennung teilweise sichtbarer Objekte in Großkopf. R E (Ed.) 12 DAGM-Symposium, Oberkochen-Aalen 1990, Bd 254 Berlin, Heidelberg, New York Springer 1990 pp 522-529
- [Tropf 80] Tropf, H. Analysis-by-Synthesis Search for Semantic Segmentation - Applied to Workpiece Recognition. Proc 5th ICPR, Miami, USA, 1980, 1980 pp 241-244
- [Winter75] Winter, H. Zenz, H P. Jenschek, J. Messung der Vertikalbewegung eines Flugzeugs. DFVLR-Forschungsbericht 75-41 1975
- [Wuest 80] Wuest, W. Pressure and Flow Measurement AGARD-AG-160-Vol 11, 1980

Subject Matter Keywords

Bord autonomes flight-test flight inspection system
inertial navigation image based navigation baroinertial altitude, inertial altitude

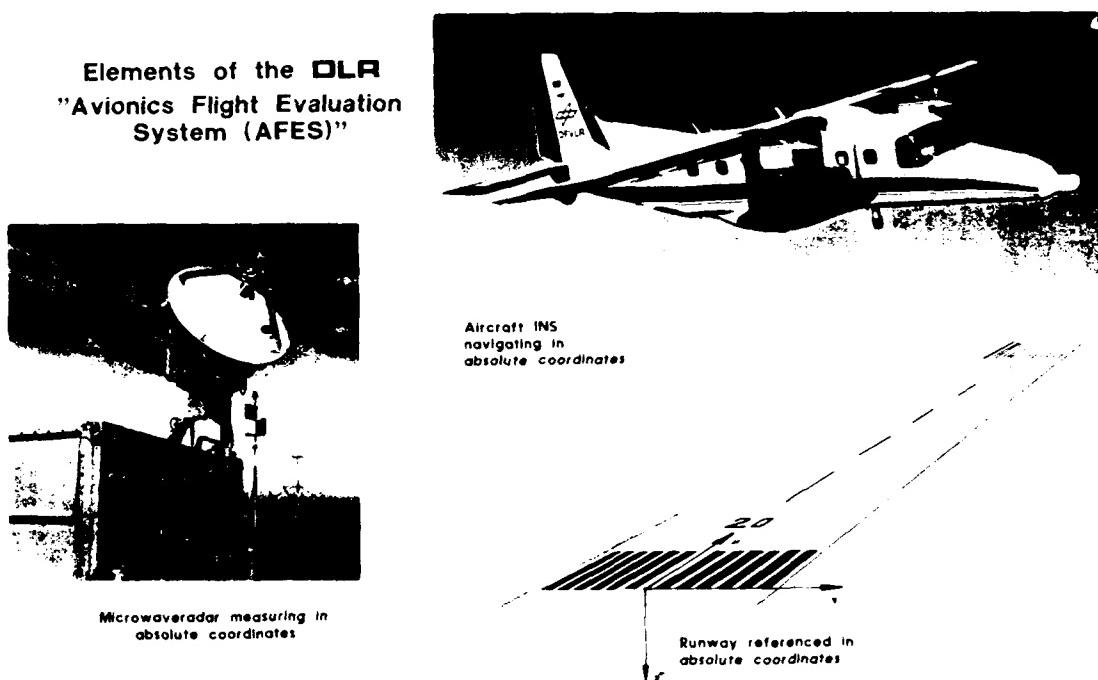


Fig. 1.1 Elements of the DLR "Avionics Flight Evaluation System (AFES)"

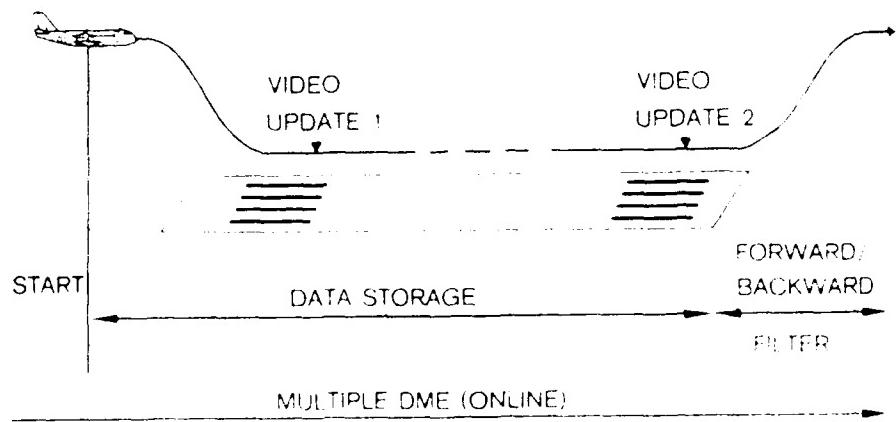
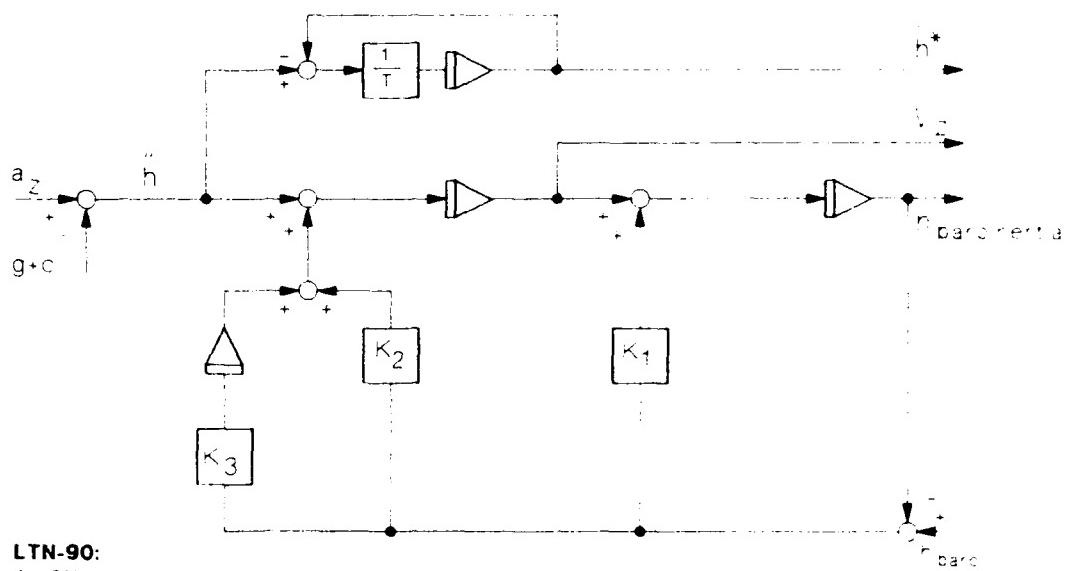


Fig. 2.1 Flight Inspection Procedure with the "Digitales Meßsystem 88 (DMS88)"

**LTN-90:**

ALIGN: $K_1 = 3/20$	NAV: $K_1 = 3/100$
$K_2 = 3/400$	$K_2 = 3/100/2$
$K_3 = 1/8000$	$K_3 = 1/100/3$

Fig. 3.1 Baroinertial Altitude Computation in an INS

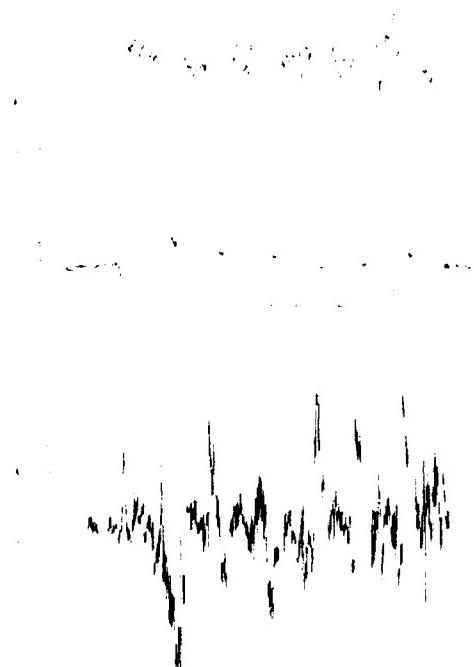


Fig. 3.2 Barometric Altitude and Altitude Error of the Twin Engine Aircraft DO 228 During Several Runway Approaches

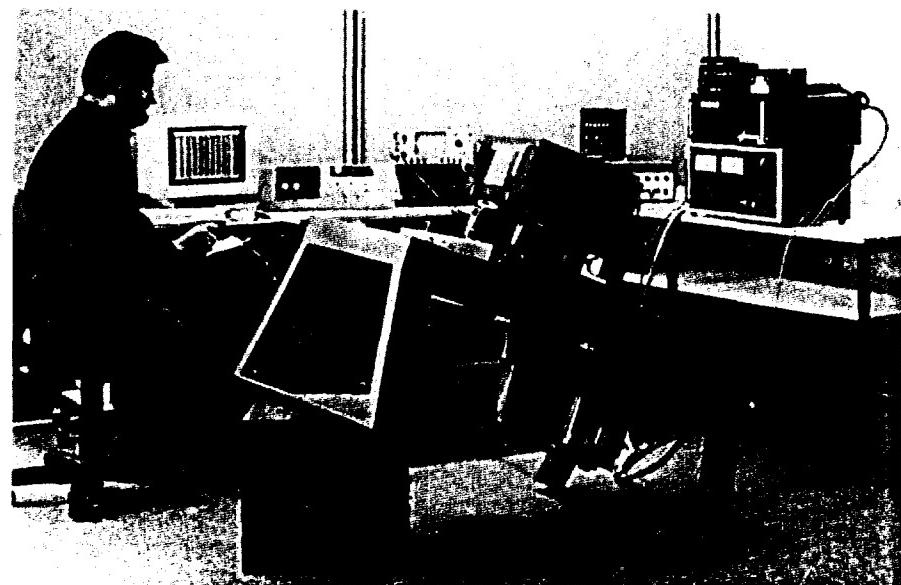


Fig. 3.3 The Litton LTN-90 INS on a Rocking Table

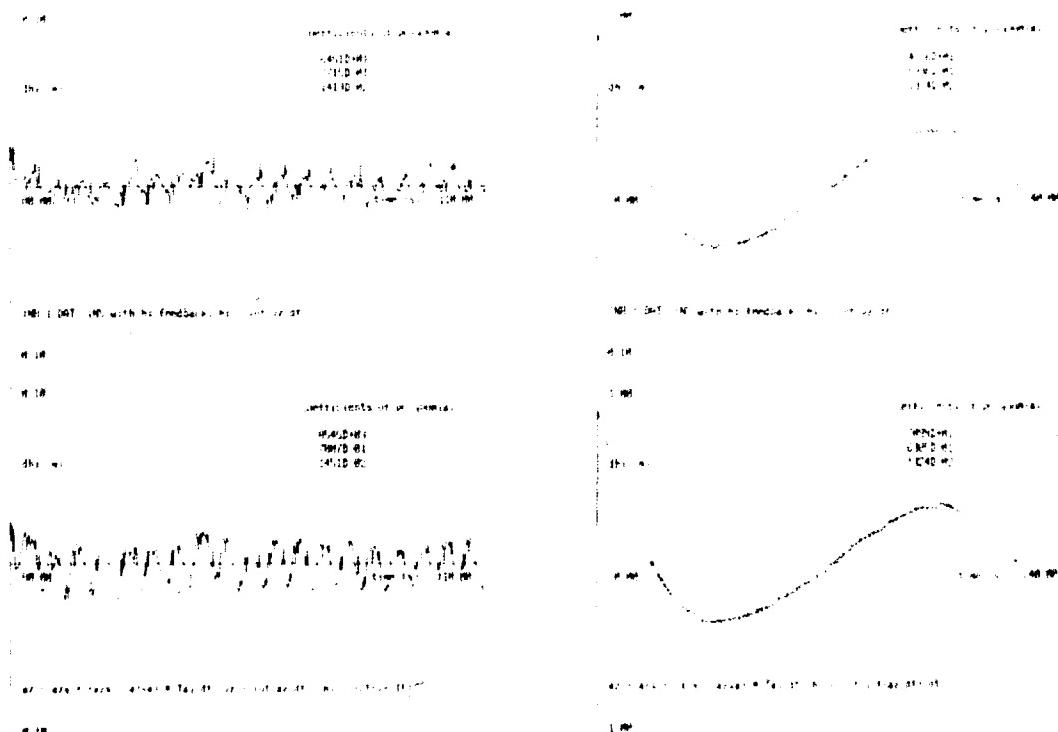


Fig. 3.4 Inertial Altitude Residues Obtained from h_1 Feedback (a) and from the Inverse Filtered Acceleration (b) with Respect to a 2nd Order Polynomial

Fig. 3.5 Inertial Altitude Residues Obtained from h_1 Feedback (a) and from the Inverse Filtered Acceleration (b) with Respect to a 2nd Order Polynomial

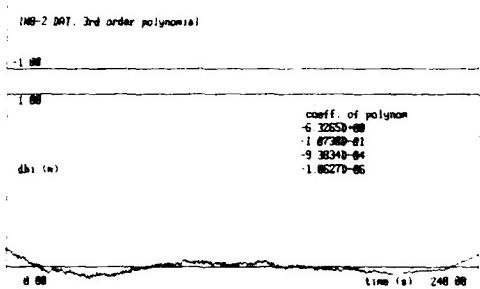
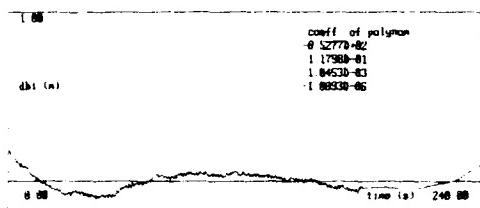
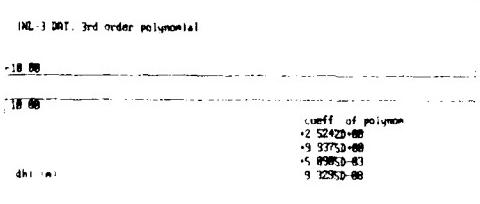
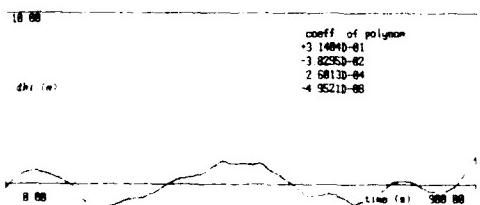


Fig. 3.6 Inertial Altitude Residues Obtained from the Inverse Filtered Acceleration with Respect to a 3rd Order Polynomial



SH187 DAT. v1=1m, v2=18m/s, v3=.881g, v4=.2m/s, g=9.81m/s², 3rd order polynomial

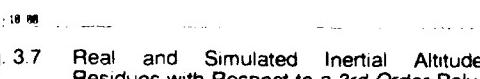


Fig. 3.7 Real and Simulated Inertial Altitude Residues with Respect to a 3rd Order Polynomial

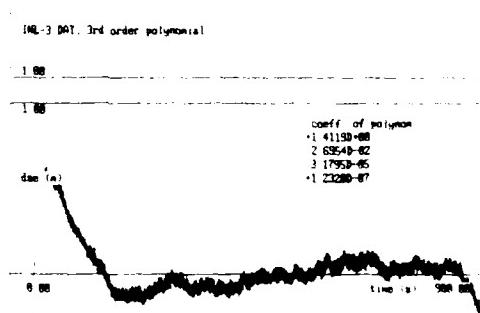
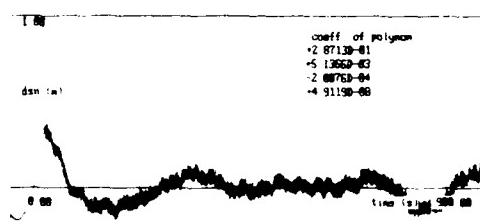


Fig. 3.8 Inertial Horizontal Position Residues from Integrating Velocities with Respect to a 3rd Order Polynomial

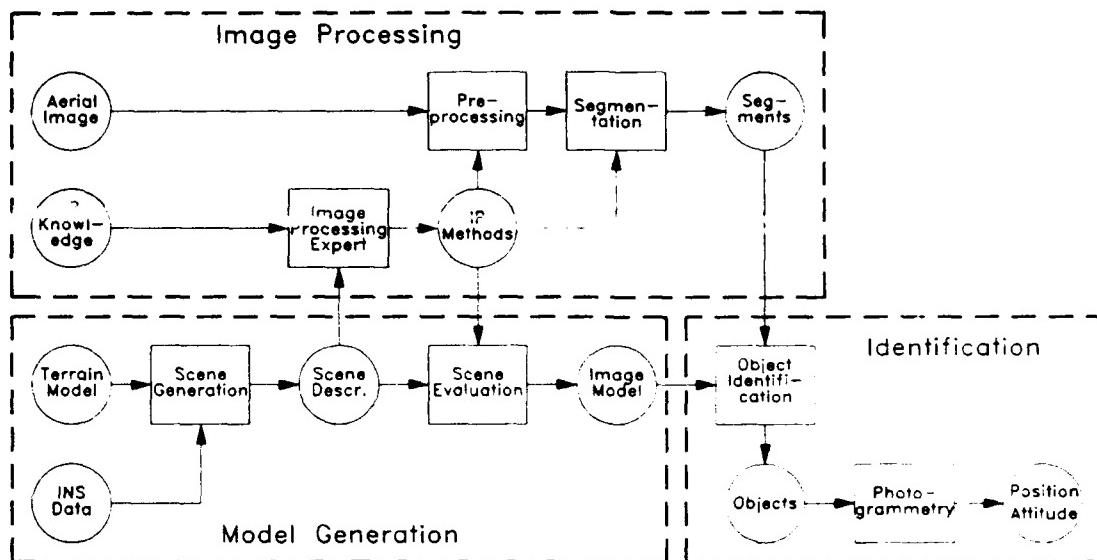


Fig. 4.1 Block Diagramm for the Scence Correlation Navigation

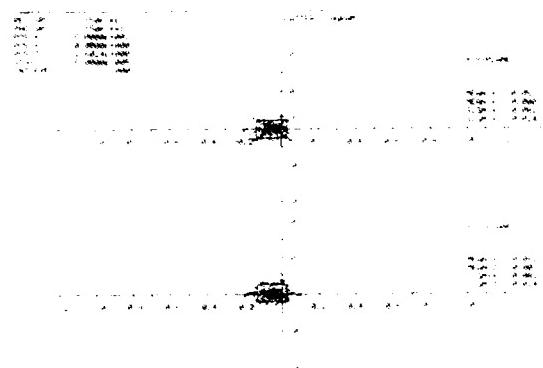
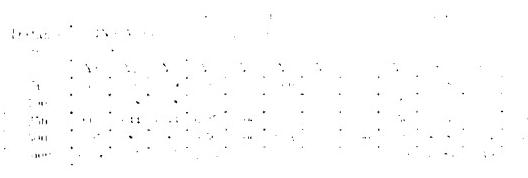


Fig. 5.1 Scatter Diagram for the Camera-Derived Position from 1000 Samples at a Distance of 50 m from the Baseline of the Runway



Tab. 5.1 Position Errors for Different Distances from the Baseline of the Runway under the Assumption that the INS Angles are Measured with an Accuracy of 0.1 Degree



Tab. 5.2 Position Errors for Different Distances from the Baseline of the Runway under the Assumption that the Aspect Angles are Estimated from the Image Data

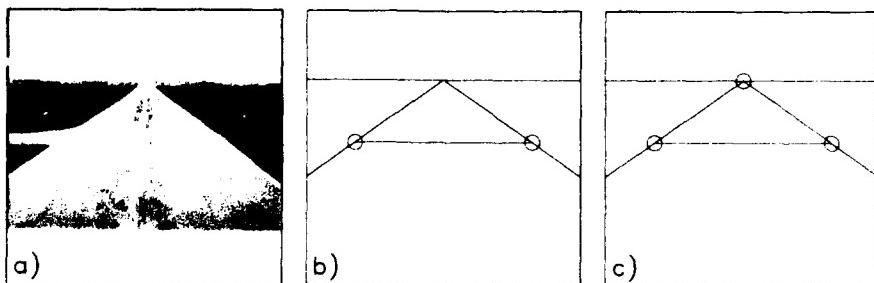


Fig. 5.2 a) Image of the Runway, b) Used Image Data for Estimating the Position only, c) Used Image Data for Estimating the Aspect Angles and the Position

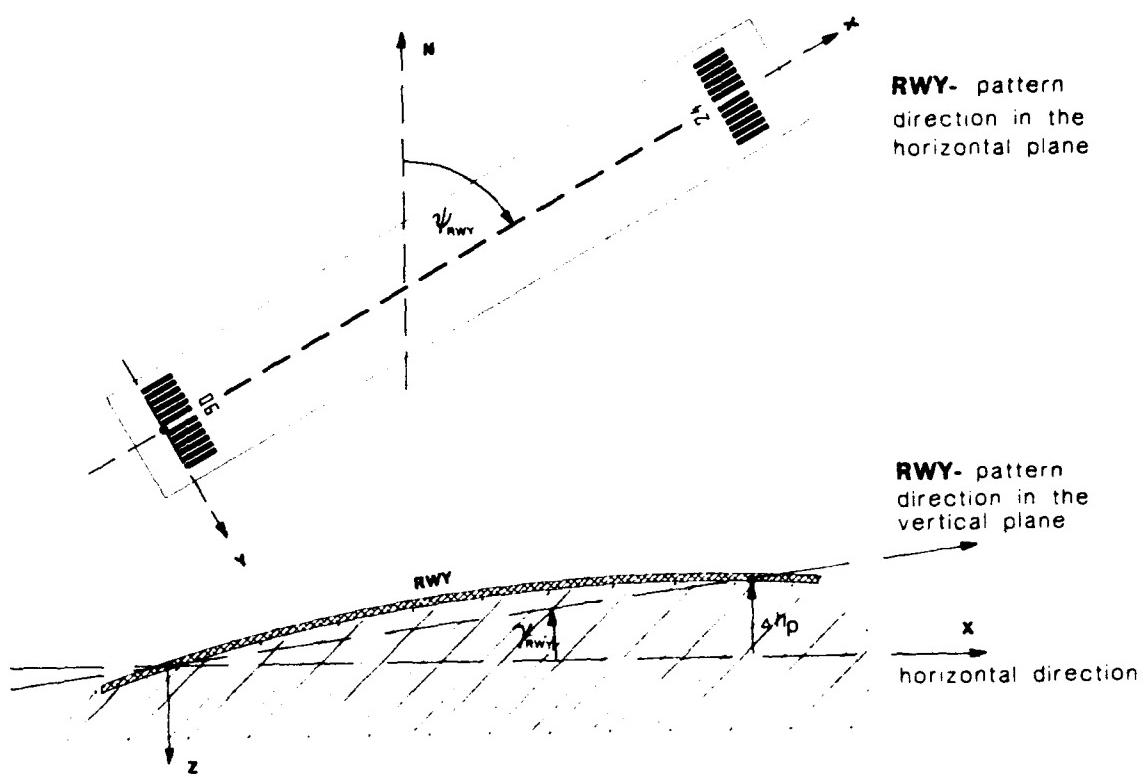


Fig. 6.1 Coordinate System Layed Down by the Runway Patterns

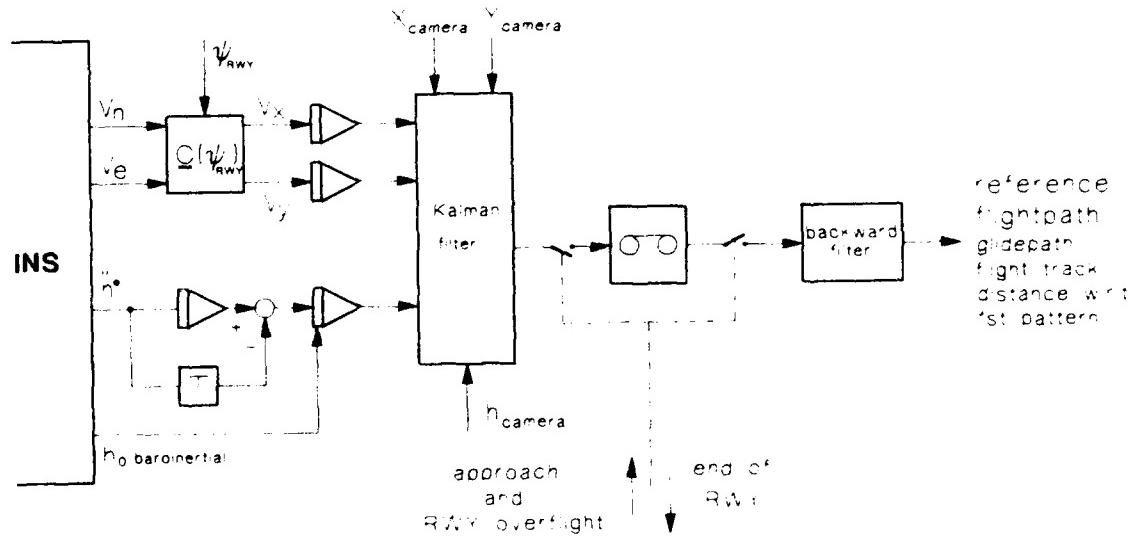


Fig. 6.2 Runway-Referenced Flight Path Computation

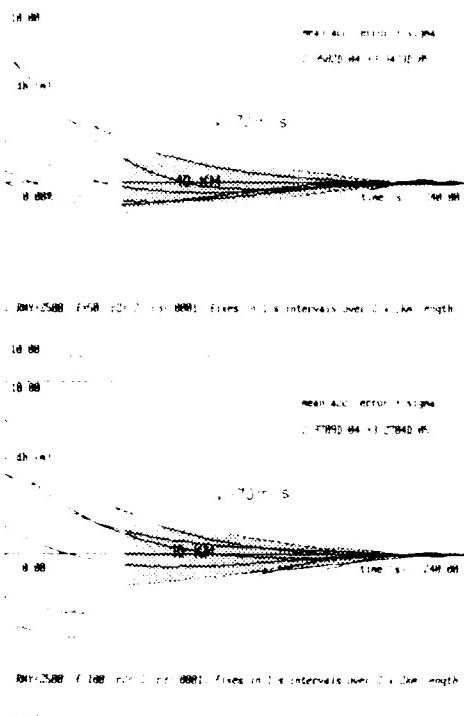


Fig. 6.3 Effect of the Camera Focal Length on the Altitude Estimation Residues for Different Runs and $\sigma(\delta h_b) = 5\text{ m}$ (shaded area = 0.01 deg limits)

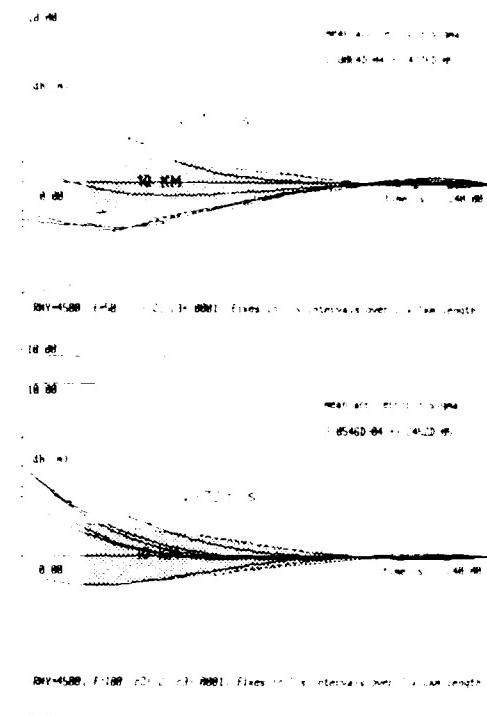


Fig. 6.4 Effect of the Camera Focal Length on the Altitude Estimation Residues for Different Runs and $\sigma(\delta h_b) = 5\text{ m}$

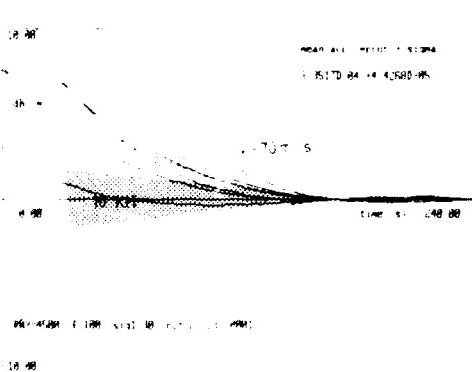


Fig. 6.5 Altitude Estimation Residues for Different Runs and $\sigma(\delta h_b) = 30\text{ m}$

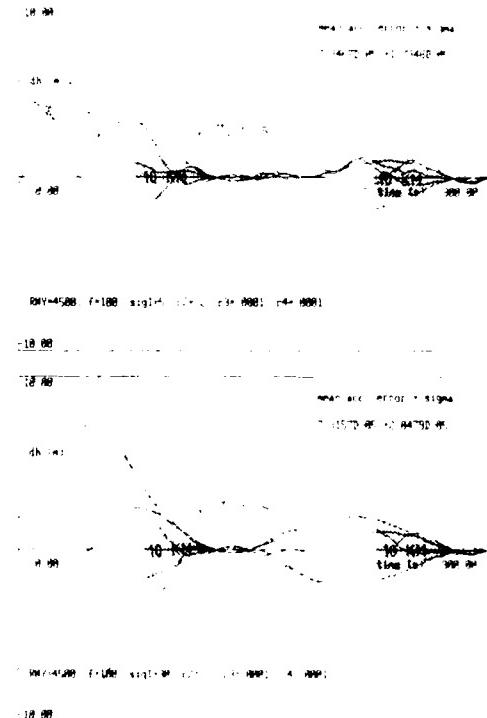


Fig. 6.6 Altitude Estimation Residues for Two Successive Runway Overflights (shaded)

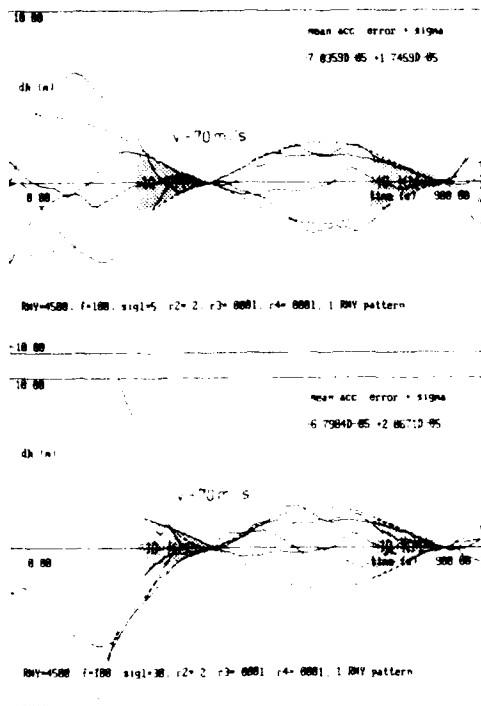


Fig. 6.7 Altitude Estimation Residues Without Knowledge of RWY Pattern Position

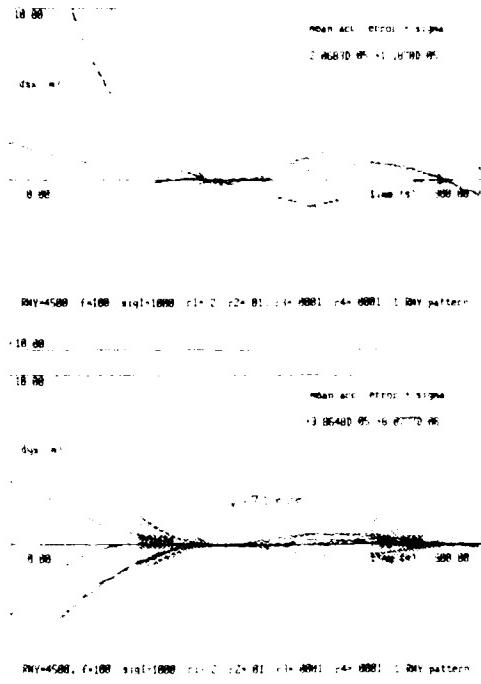


Fig. 6.8 Horizontal Position Estimation Residues Without Knowledge of RWY Pattern Position

An Hierarchic Alliance of Filters for Fault Tolerant Navigation

Using Two Inertial Systems with Aiding Sensors

J. Chris McMillan, Jeff S. Bird
 Communication and Navigation Section
 Defence Research Establishment Ottawa
 Ottawa, Ontario, Canada K1A 0Z4

Dale A.G. Arden
 Computing Concepts Ltd.
 10 Phylis St.
 Nepean, Ontario, Canada K2J 1V2

1. SUMMARY

A Dual Inertial Integrated Navigation System (DIINS) is being developed for the Canadian Navy to improve the navigational accuracy and reliability on ships which have two inertial navigators plus other aiding navigation systems and sensors such as GPS, Loran-C, Omega, Doppler Speed Log(s) and so on. The sensor integration architecture being proposed to optimally combine all navigation sensors on such a vessel is an "hierarchic alliance" of Kalman filters, which is designed to allow sensitive error compensation as well as complete fault detection isolation and reconfiguration (FDIR). This architecture is ideally suited to central processing, can take advantage of parallel processing, and provides significant advantages over both the conventional unifilter approach and the "federated" (or cascaded) filter approach.

This hierarchic alliance consists of a specific set of optimal filters running in parallel, with each filter processing measurements from a different subset of the navigation sensors. These filters can be partially ordered so that primary and secondary filters can be defined. The primary filter(s) provide the optimal navigation solution, while the secondary filters provide uncorrupted backup in the event of a sensor fault.

The primary motivating factor for this architecture is to provide optimal integration under all conditions, and in particular after the occurrence of subtle sensor faults which could not be immediately detected and which could therefore corrupt the primary filter(s). This alliance of filters can provide an uncorrupted optimal solution, since it can be configured so that there will always be a secondary filter which, at the

time of failure, was running independently of the faulty sensor. This removes the usual need to "back out" of a failure which was not immediately detected and thus substantially simplifies reconfiguration in response to such a failure.

Another motivating factor for this architecture is that the partial independence of the parallel filters also facilitates the detection and isolation of sensor faults. This can be accomplished by multiple levels of statistical hypothesis testing on a set of parallel Kalman filters. Fault detection techniques used include the usual sensor data reasonableness and filter residual tests as well as a chi-square hypothesis testing technique applied to the state vectors, and inter-filter voting applied to the residuals test results.

While this approach is computationally intensive, modern software techniques, and soon to be available processing power, are expected to make the real time implementation of this hierarchic alliance of filters quite practical. The system envisioned in this paper is being designed and built at the Defence Research Establishment Ottawa for the Canadian Navy and is expected to see initial real-time sea trials in 1993.

2. INTRODUCTION

During the 1980's the Defence Research Establishment Ottawa (DREO) developed a Kalman filter based Marine Integrated Navigation System (MINS), as described in McMillan (1990), that optimally integrated the navigation sensors found aboard Canadian naval vessels at the time. These included GPS, Transit, Loran-C, Omega, speed log, gyrocompass, and operator entered sextant measurements and position fixes. MINS employs a

single 17-state, dynamically reconfigurable, complementary Kalman filter (UD formulation) with the "system" states representing the dead reckoning errors. MINS units are being installed on all major Canadian naval vessels with the exception of the new patrol frigates which are still under construction. Unlike existing ships, these new frigates do not have dead reckoning systems, but are being equipped with twin inertial navigation systems (installed fore and aft) rather than a gyrocompass. Future submarines are expected to be similarly equipped, and present the added problem of submerged navigation, which is largely unaided. Consequently, a next generation of the MINS system will be required to take full advantage of the redundancy available from the INS's on the new vessels.

The new integrated system currently being designed and simulated at DREO is called DIINS (Dual Inertial Integrated Navigation System) and is a substantial revision of MINS. It will optimally integrate at least the following sensors:

- INS_1 ,
- INS_2 ,
- GPS,
- Loran-C,
- Omega,
- speed log,
- Hyperfix,
- operator entered sextant fixes.

The presence of two inertial systems has required a completely new set of error models in the Kalman filters and their redundancy allows a much higher degree of fault tolerance to be built into the system. The primary objective of DIINS is to enhance the accuracy and reliability of the navigation system, and the FDIR architecture and algorithms are the key to accomplishing this. The algorithms are designed to detect and isolate (identify) sensor failures promptly and to automatically remove the failed sensor from the navigation solution without introducing a significant discontinuity in the output. Since full failure mode operation requires that DIINS be designed to integrate any viable subset of the marine sensors mentioned above, this flexibility provides the added benefit that DIINS could be used on virtually any marine platform.

The primary reason for having two relatively expensive INS's on each ship is to provide reliability in the event of a failure. Much effort has therefore been directed towards the prompt detection and isolation of subtle faults in an INS, such as an accelerometer bias shift (perhaps due to temperature control failure) or an out of spec. gyro bias shift. It will be shown that this can be accomplished, even in the submarine scenario where the only aiding sensor is the speed log.

Without the sensor integration provided by DIINS, a slow failure of an INS could not be quickly detected. It would not be clear that an INS has failed until the discrepancy becomes quite large. Early failure isolation (determining which INS has failed), would be even more difficult, especially in the case of a submerged submarine, where accurate aiding sensor position measurements are not available.

This paper outlines an optimal filtering architectural concept, and how it can be applied to this particular set of sensors. It also briefly describes the basic Kalman filter design, and various techniques currently being considered and implemented for fault detection, isolation and reconfiguration (FDIR) in this context. These include:

- measurement reasonableness tests,
- Kalman filter residual tests,
- multiple filter residual test voting, and
- a χ^2 (chi-square) filter self test technique that is applied to the state estimate and involves propagating a "shadow" filter (with no measurements) for each regular Kalman filter.

Some simulation results are also presented to illustrate the power of the FDIR techniques to detect and isolate subtle inertial sensor faults, even in the submarine case where only speed log aiding is available.

Since DIINS is still at an early stage of development, the designs outlined in this paper are only preliminary. Full simulations have yet to be completed to verify the performance of the fault detection, isolation and reconfiguration algorithms under a full set of expected failure conditions.

3. HIERARCHIC ALLIANCE OF FILTERS

3.1. General Concepts

To discuss the proposed architectural concept, some notation will be helpful. Consider the usual case where one continuous system (inertial or dead reckoning) is to be integrated with n aiding sensors (GPS, Loran-C etc.). The usual unifilter shall be referred to as F^0 . This is envisioned as a complementary or error state filter, with a state vector partitioned as follows:

X_0	inertial system / sensor error states	(1)
X_1	first aiding sensor error states	
X_2	second aiding sensor error states	
X_n	n^{th} aiding sensor error states	

The measurement vector can also be decomposed:

Z_1	first aiding sensor measurement
Z_2	second aiding sensor measurement
Z_n	n^{th} aiding sensor measurement

Let F^i (for $i = 1, 2, \dots, n$) refer to the filter in which the i^{th} aiding sensor states and measurements have been removed. Let F^{ij} refer to the filter in which the i^{th} and j^{th} aiding sensor states and measurements have been removed. Similarly define F^{ijk} and so on.

The F^1 can be considered the first generation filters, the F^2 second generation, and so on. The set of all such filters form a tree, as shown in Figure 1. Each filter in this tree can easily be spawned from the filter immediately above it (its parent filter), by simply eliminating the appropriate states and measurements. This reconfiguration can quite easily be done dynamically (in real time) without introducing any discontinuity, since the remaining states will be unchanged, as will their covariance matrix.

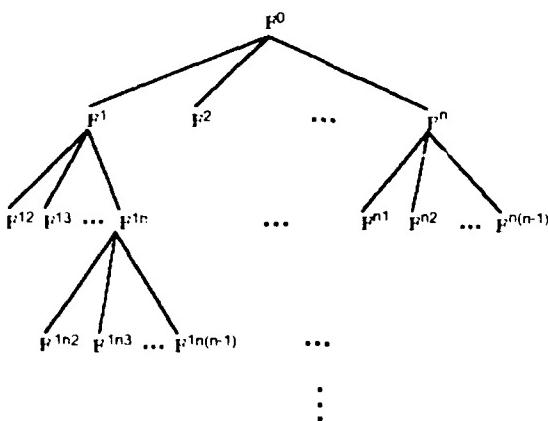


Figure 1. Hierarchic Alliance Structure

Any set of filters sharing a common parent filter, together with that parent shall be referred to as a family of filters. In the proposed approach, only one such family of filters is required at any given time. This provides the optimal filter (the parent) and a set of backup filters. The fact that each backup filter is independent of one of the sensors assists in all three aspects of fault detection, isolation and reconfiguration.

This architecture provides many advantages over both the conventional unifilter approach and the "federated" (or cascaded) filter approach. The different filters share sensor measurements only, rather than state vector information and therefore can be designed and

run completely independently. They are therefore not subject to the filter design constraints which lead to sub optimality of the federated filters.

As discussed below, the partial independence of the parallel filters substantially improves FDIR, with failure detection being more sensitive, isolation more decisive, and reconfiguration much simpler.

With one INS and n aiding sensors it would be necessary to run at most $n+1$ filters at a time. With practical considerations, this number may be reduced. Although this may be considered burdensome, with modern software techniques and recently available processing power, the practical implementation of this hierarchic alliance of filters is now possible.

3.2. Reconfiguration

This architecture provides substantial capability for reconfiguration in the event of sensor failure, which is not available to a unifilter or to federated filters. The greatest difficulty with reconfiguration under both the unifilter and the federated filter schemes, is that the primary filter may have been corrupted by the sensor failure. This requires either a complete filter reinitialization, with the resulting total loss of state estimate information, or some complex "backing out" scheme which attempts to use the state estimate from some pre-selected point in the past, hopefully before the failure occurred. The hierarchic alliance approach resolves this problem by simply having "hot" backup filters running at all times.

For any given set of aiding sensors, there will be a maximal filter in the complete tree (see Figure 1) which integrates all available (not failed) sensors. To provide an uncorrupted filter for single failures, it is only necessary to run the family containing the maximal filter as parent. For example if F^i is the maximal filter, and a soft failure (one who's detection is delayed) occurs in sensor j , then F^i will have been corrupted, but F^{ij} will not have been. Therefore F^{ij} can be used as the new maximal filter, with new subfilters F^{ijk} being spawned from F^{ij} . This is the basic hierachic alliance reconfiguration concept.

It should be mentioned that if the above failure in sensor j was a hard failure, then F^j will not have been corrupted, in which case F^j would be used to spawn F^{jk} directly, since F^j should have better state estimates than F^k . Therefore if sensor j is not expected to ever have soft failures then the corresponding filter F^j is not necessary for reconfiguration purposes. However, these secondary filters have an equally important role in fault isolation, as described below.

It should also be mentioned that inherent in this approach is the assumption that a second soft failure

will not occur in the time interval between the occurrence and detection of the first. It is therefore important to be able to detect these soft failures as quickly as possible in order to minimize this possibility.

Extending this technique to handle multiple simultaneous failures is possible and in fact is conceptually quite simple. To handle m simultaneous failures, it would be necessary to include all different filters for m generations below the maximal filter in the tree of Figure 1. For reasons of practicality however this has not been pursued further.

3.3. Failure Isolation

This architecture provides significant fault isolation capability which is not available to a unifilter or to cascaded filters.

When a sensor failure occurs, one consequence is that the error model assumptions underlying the Kalman filter are no longer valid. This can lead to unanticipated effects. Therefore some of the most sensitive fault detection schemes indicate only that there is a statistically significant problem within a Kalman filter but give no indication of where the fault lies.

For example the chi-square test indicates only that there is a statistically significant discrepancy between a Kalman filter's design model (represented by covariance information) and the measurements being processed by that filter (represented by state estimate information). Assuming that the fault is not with the filter design, it can then only be deduced that the sensor data is behaving in unexpected way, most likely due to a sensor failure. Some other means must then be used to determine which sensor is at fault. In an hierachic alliance family of filters however, one and only one of the filters should be unaffected by the failure of a single sensor, namely the filter not using the failed sensor. It should therefore be possible to apply simple deductive logic to isolate the failed sensor. This is described in more detail in section 5.3 below, in the context of DIINS.

Furthermore, some standard FDI techniques, such as residual testing, are often used to isolate failures within a single filter even though they are actually ambiguous in this context. As will be explained in section 5.2 below, this ambiguity can be resolved by using residual test results from the multiple filters proposed.

3.4. Failure Detection

Failure detection and isolation are not entirely separable, since techniques which can detect but not isolate a failure are of limited use. The enhanced fault

isolation capability of this architecture, as described above, therefore improves fault detection. It does this by allowing the use of more sensitive detection schemes such as the chi-square test, which would not otherwise allow isolation.

Ultimately sensor failures are detected by comparisons made between different sensor measurements (using *a priori* information regarding their different error behaviours and perhaps also regarding absolute physical limits on platform dynamics). A full family of filters, as described above, provides a richer set of sensor comparisons than is available to a unifilter or to a federation of filters. This should also improve the likelihood of failure detection.

With several different filters providing somewhat independent test results, voting can be used. This provides a level of verification not available to a unifilter, or even to a federation of filters. This can be used in two ways, or some combination of both:

- individual failure detection trip levels can be made lower (more sensitive) while maintaining the same level of confidence, or
- trip levels can be left the same, so that detection confidence would increase (false alarms reduced).

4. DIINS FILTERS

The presence of the second INS in DIINS slightly complicates this picture, since there will initially be two trees such as shown in Figure 1, plus a special purpose INS1/INS2 filter. Since both trees are equivalent, and are derived from the same unifilter, a brief description of this unifilter will be given.

4.1. The DIINS Unifilter

The usual multisensor filtering approach is taken. For navigation purposes this is normally a complementary, or error state, filter. This type of filter primarily estimates the errors of a continuous system such as an inertial (or dead reckoning) system, and coincidentally estimates the aiding sensor errors that are observable. The navigation output then comes from the primary sensor, corrected by the appropriate state estimates.

To minimize the effect of non linearity in the measurement equation, an extended filter formulation is used, whereby the filter measurements are formed by taking the difference between aiding measurements (such as GPS pseudo ranges etc.) and a prediction of what the aiding measurement should be at the filter's estimated position (or velocity etc.).

As long as an INS is functional, it is considered the primary sensor for position, velocity and attitude for navigation and related command and control

functions. In this case the INS error estimates are used to correct the output of the INS for these purposes.

The details of the error modeling developed for the unifilter are well beyond the scope of this paper, however it is useful to display the state vector and measurement vector. The unifilter state vector for DIINS is shown in Table 1.

Table 1. DIINS Unifilter State Vector

State Number	Description
1-2	INS horizontal position error
3-4	INS horizontal velocity error
5-7	INS platform misalignment
8-9	horizontal accelerometer biases
10-12	gyro biases
13-18	GPS Pseudo range errors
19	GPS clock bias
20	GPS clock drift
21-22	Speed Log errors
23-24	Ocean Current
25-26	Loran-C hyperbolic errors
27-33	Omega phase errors

This has been partitioned into the inertial system states (9), the inertial sensor error states (6), the GPS states (8), the Speed Log error states (4), the Loran-C error states (2) and the Omega error states (7). Of course there are two such filters corresponding to the two inertial systems.

The measurement vector can be similarly partitioned, as shown in Table 2. It is expected that more sensors may be added to this list as development proceeds.

Table 2. DIINS Unifilter Measurement Vector

Measurement Number	Description
1-2	Speed Log
3-8	GPS pseudo ranges
9-14	GPS pseudo range rates
15-16	Loran C time delays
17-23	Omega phase differences
24-25	position fix

4.2. DIINS Filter Configuration

Since the two INS's provide more than one primary sensor, in the context of complementary filtering, it was only natural to consider multiple filters, even though it is still possible to construct a complementary unifilter. Although such a unifilter would provide optimal integration under normal conditions, a gradual sensor failure could corrupt the entire state vector before it was detected, causing serious problems. Individual filters that would combine one INS and one additional sensor, such as the federated filter of Carson (1987), were considered. This consists of a filter for each INS - single aiding sensor combination and one "federal" filter to integrate the subfilters. However, the interdependency and sub optimality of the filters, and the reliance on an uncorrected federal filter precludes the availability of an optimal backup filter, and thus does not provide completely graceful degradation.

It was therefore decided to investigate an architecture such as described in section 3 above. This expands on ideas drawn from Widnall (1987).

Another basic characteristic of the DIINS system is a special filter called the INS1-INS2 filter that is very useful in helping to detect and isolate INS failures. Since this filter has very low measurement noise, it will be very sensitive to INS failures. However it is not directly used in the navigation solution since it estimates only the relative errors between the two INS's.

If the aiding sensors are referred to as S1, S2, S3 and S4 (representing for example GPS, Loran-C, Speed Log and Omega), then the initial set of filters comprising the hierachic alliance are listed in Table 3, with the alliance structure illustrated in Figure 2. After reconfiguration this set will of course change, and if one of the INS's fails, then there will only be one tree. If both INS's fail, then a non-complementary filter would be required to integrate the remaining sensors. This level of reconfiguration is beyond the scope of this paper.

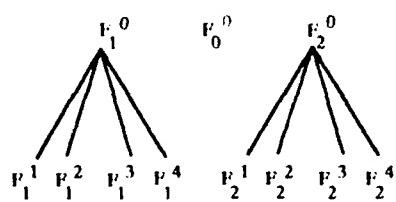


Figure 2. DIINS Initial Set of Filters.

Table 3: Initial Set of Filters for DIINS

Filter	SENSORS
F_0^0	INS1, INS2
F_1^0	INS1, S1, S2, S3, S4 (INS1 unifilter)
F_2^0	INS2, S1, S2, S3, S4 (INS2 unifilter)
F_1^4	INS1, S1, S2, S3 (i.e. F_1^0 w/o S4)
F_2^4	INS2, S1, S2, S3 (i.e. F_2^0 w/o S4)
F_1^3	INS1, S1, S2, S4 (i.e. F_1^0 w/o S3)
F_2^3	INS2, S1, S2, S4 (i.e. F_2^0 w/o S3)
F_1^2	INS1, S1, S3, S4 (i.e. F_1^0 w/o S2)
F_2^2	INS2, S1, S3, S4 (i.e. F_2^0 w/o S2)
F_1^1	INS1, S2, S3, S4 (i.e. F_1^0 w/o S1)
F_2^1	INS2, S2, S3, S4 (i.e. F_2^0 w/o S1)

5. DIINS FAILURE DETECTION

Of course, sometimes fault detection is trivial to the navigation officer monitoring the sensors. The DIINS system must be aware of such obvious faults even while it is monitoring the integration filters for subtle faults that the operator has yet to see. This suggests multiple levels of fault detection to be built into the system that can roughly be ordered as follows:

1. Operator disable - Often the operator knows that a sensor is faulty and should not be integrated.
2. Built-in-Test - The sensors' built-in-test indicators are used at a very high level to isolate a faulty sensor.
3. Reasonable Data - The raw sensor data is subjected to upper and lower bounds to ensure that it is physically realistic.
4. Consistent Data - These are tests on the physical reasonableness of the *change* in sensor output from one measurement to the next.
5. Reasonable States - The state estimates are tested against physically meaningful upper and lower bounds.
6. Residual Tests - The integration filter compares the magnitude of the filter residuals associated with the incoming sensor data with what it expects them to be based on the error models in the filters.
7. Inter-Filter Residual Test Voting - Simple deductive reasoning is applied to the results of the

individual filter residual tests to isolate failed sensors.

8. Filter Self Test - Each filter can test its own state vector against the state vector of a "shadow" filter (an identical one propagated with no measurements).

The first 5 of these tests provide a very efficient and easily implemented prescreening of sensor measurements, before the more sensitive but more complex tests are performed. These simple tests are very effective in protecting filters from spurious measurements and rapid sensor failures.

The sixth and seventh (residual testing) provide a more sensitive, statistically based test which is effective in detecting and isolating discontinuity-type failures (hard failures) and is described in the following section. The eighth test (chi-square) is another statistically based test, which is most sensitive in detecting and isolating slow accumulative-type failures (soft failures). This is also described in more detail below.

5.1. Residual Testing

Assuming that a sensor measurement has passed the preliminary FDI tests (the first 5 tests listed above), it must still pass the residual test before being incorporated into the filter. A brief discussion of filter mechanization is needed to explain this test.

The filters are mechanized with Bierman's UDU^T form with scalar measurement updating and diagonalized noise covariance matrices. If $z(k)$ denotes the measurement vector for a filter at time k then:

$$z(k) = Hx(k) + w(k) \quad (3)$$

where H , $x(k)$, and $w(k)$ are the corresponding measurement geometry matrix, state vector and measurement noise vector, respectively. In a complementary error state filter, such as used in DIINS, these measurements z are generally *miscreations*, or differences between some quantity y_{meas} measured by an aiding sensor and a prediction of what that quantity should be, based on the position, velocity etc. of the primary sensor (the INS) y_{INS} .

$$z(k) = y_{\text{meas}}(k) - y_{\text{INS}}(k) \quad (4)$$

The component i of the residual vector v , is then defined as

$$\nu_i(k) = z_i(k) - E[z_i(k)|\dot{x}(k|k-1)] \quad (5)$$

where $\dot{x}(k|k-1)$ is the estimate of the state vector of the filter at time k given all data up to time $k-1$. If the covariances of the state vector error and the measurement noise vector are denoted by P and R , respectively, then the variance of the residual is

$$E[\nu(k)\nu^T(k)] = HP(k|k-1)H^T \cdot R(k) \quad (6)$$

where we have assumed a linear measurement equation (Eq. 3) and no correlation between the state vector error and the measurement noise.

If neither INS has failed, the DIINS residual testing algorithm uses F_0^0 and the most accurate pair of filters remaining (F_1^0 and F_2^0 in the no fail case). Note that each sensor is used in 2 of the 3 filters. Sensor failure detection and isolation using residual testing is implemented as follows:

1. At each update time, form the residual for each measurement in each filter, Eq. 5.
2. Form the estimate of the variance of each residual from the H , P , and R matrices from each filter according to Eq. 6.
3. If the square of the residual exceeds some preset number, A^2 , times its filter computed expected variance:

$$\nu_i^2(k) > A^2[HP(k|k-1)H^T \cdot R(k)]_{ii} \quad (7)$$

then that particular measurement is considered bad and is not incorporated in the Kalman filter update (typically A is set to 3 to provide a "3-sigma" test).

This is the usual application of residual testing within a single complementary filter, and constitute FDI test level 6 of the previous section. From Eq. 4 and Eq. 5 it can be seen that two assumptions are implicit in this method:

- A. the residual ν in Eq. 5 is large because of the misclosure z rather than the state estimate $\dot{x}(k|k-1)$, and
- B. the misclosure z in Eq. 4 is large because of the aiding sensor measurement y_{meas} rather than the primary sensor prediction y_{INS} .

Within a single complementary filter these are reasonable hypotheses because there is no useful response under the alternative hypothesis (a filter failure or a primary sensor failure). With the proposed architecture however, there is both the means to determine where the fault lies and the means to respond appropriately if the filter or the primary sensor (the INS) has failed. This is especially true when two primary sensors (INS's) are available, as is the case with DIINS.

5.2. Inter-Filter Residual Test Voting

The usual residual testing, described by the three steps above, can be extended as follows, to properly deal with failure of the primary sensor:

4. If a filter residual fails the 3-sigma test for at least M of the last N updates, then that residual is declared "suspect" in that filter.
5. If a residual is declared "suspect", then both sensors associated with that residual (the aiding sensor and the primary sensor) are declared suspect.
6. If a sensor i (aiding or INS) is "suspect" in all top level filters which use it, and its residual-associated sensor (INS or aiding) is not suspect in at least one top level filter which uses it, then the sensor i is declared failed.

Steps 5 and 6 above describe roughly how to apply deductive logic to these residual test results in order to isolate the failed sensor. Step 6 eliminates assumption B above by ascribing failure to the appropriate sensor. Assumption A is arguably less important, however it too can be somewhat eliminated by adding to step 6 a requirement intended to verify the validity of the state vector estimate. This is more difficult to couch in general terms, but in the case of DIINS it can be assumed that if the residual associated with the most accurate aiding sensor (and the INS) is not suspect, then the state estimate is not suspect.

In the DIINS baseline system of two INS's and four secondary sensors, the residual tests would be conducted in all 11 filters and measurements would only be included in those filters if they passed the residual tests. The top level filters, as shown in Figure 2 above, are the two unifilters, F_1^0 and F_2^0 , along with the INS1-INS2 filter, F_0^0 . Only these would participate in the sensor fault isolation of Steps 5 and 6 above. Table 4 below shows the logic table implementing these steps. This table is not complete, however it does include all decisive cases.

Table 4: Residual Test FDI Voting

F_1^0 Residuals (INS1/S1/S2/S3/S4)				F_0^0 INS1 INS2				F_2^0 Residuals (INS2/S1/S2/S3/S4)				Failed Sensor
S4	S3	S2	S1		S1	S2	S3	S4				
.	.	X	X		0	.	.	.	INS1			
.	.	0	X		X	.	.	.	INS2			
.	0	X	0		X	0	.	.	S1			
.	X	0	0		0	X	.	.	S2			
.	X	0	0		0	.	X	.	S3			
X	.	0	0		0	.	.	X	S4			

0 - Residual is not suspect

X - Residual is suspect: M out of N residual rejections

- - Irrelevant

An "X" in the table means that a residual associated with that aiding sensor has failed the "3-sigma" test for M out of the last N samples, so that the residual is "suspect." For the aiding sensor to be declared failed, a residual formed from that sensor in both unifilters must be suspect and there must be evidence that neither the primary sensor (INS) nor the state vector estimate are at fault.

Acceptable residuals in F_0^0 provide evidence that the primary sensor (INS) is not at fault, while acceptable residuals for the most accurate remaining sensor in the unifilters provide evidence that the unifilter state vectors are not at fault.

5.3. Filter Self Test - The Chi-Square Test

Not all failures will be observable through filter residual monitoring. Any slowly accumulating error, such as an unmodeled INS errors, for example, would tend to be absorbed in the state estimates of other INS errors and would show little effect in the residuals. Thus a test on the state vector of a filter as a whole, as outlined by Brumback and Srinath (1987) and based on earlier work (see Kerr (1987)), is being investigated for use in DHNS. The idea is to have two solutions of each Kalman filter run in parallel, one with measurements and the other without. (We call the filter without measurements a "shadow" filter and look for a difference between a filter and its "shadow.") This shadow filter provides for a statistically significant reasonableness test that is applied to each filter's state estimate.

Consider one filter, to simplify the explanation. Define $\hat{\delta}x$ as the difference in the two state vectors:

$$\begin{aligned}\hat{\delta}x &\equiv \hat{x}_m - \hat{x}_e \\ &= (x - \hat{x}_m) - (x - \hat{x}_e) \\ &= \hat{x}_e - \hat{x}_m \\ &= \hat{\delta}x\end{aligned}\quad (8)$$

where \hat{x}_m is the state vector estimated using measurements and \hat{x}_e is the state vector estimated without measurements (thus if \hat{x}_e is initially zero then it will always remain zero). The covariance of the difference in the state estimates can be formed from the covariances of their errors:

$$\begin{aligned}P_{\hat{\delta}x} &= E[\hat{\delta}x\hat{\delta}x^T] = E[\hat{\delta}x\hat{\delta}x^T] \\ &= E[\hat{x}_e\hat{x}_e^T] - E[\hat{x}_e\hat{x}_m^T] - E[\hat{x}_m\hat{x}_e^T] + E[\hat{x}_m\hat{x}_m^T] \\ &= P_e - P_{em} - P_{me} + P_m\end{aligned}\quad (9)$$

It is fairly easy to show that under conditions of optimal filter gains, identical state models and identical initial conditions for both solutions that

$$P_{em} = P_{me} = P_m \quad (10)$$

so that

$$P_{\hat{\delta}x} = P_e - P_m \quad (11)$$

Since $\hat{\delta}x$ is Gaussian (it being the difference of two Gaussian vectors) of zero mean and covariance given by Eq. 10, its distribution is completely defined. The test for a failure consists of computing the scalar test statistic

$$k^2 = \hat{\delta}x^T [P_{\hat{\delta}x}]^{-1} \hat{\delta}x \quad (12)$$

The test statistic k^2 has a chi-square distribution with n degrees of freedom, n being the dimension of the state vector under test. A failure is declared with a confidence level of $(1-\alpha)$ when the test statistic exceeds the appropriate threshold:

$$k^2 > \chi_{n,\alpha}^2 \quad (13)$$

where $\chi_{n,\alpha}^2$ is determined from the tables of the chi-square distribution. Note that α is the probability of false alarm; declaring a failure when there is none.

Note that this test can be just as easily applied to any

subset of the vector $\delta\dot{x}$, by simply extracting its covariance matrix from $P_{\delta\dot{x}}$. The dimension n of the subvector used will of course affect the failure threshold. One current topic of investigation is to determine the most effective set of states for the purpose of detecting gyro and accelerometer failures.

Failure of this filter self-test, nominally called the chi-square test, indicates that the Kalman filter solution has deviated from its predetermined model (under the assumption that the filter has been properly modeled and tuned). The alternative solution from the shadow filter only propagates the initial conditions using the given model and cannot be corrupted by bad data. The failure of the test is attributed to measurement errors which in turn point to undetected sensor failures. There is not enough information in the test to conclude which sensor in the filter is not performing correctly. Sensor failure isolation requires multiple filters using different combinations of the available sensors.

Consider the full set of 11 filters listed in Table 3 and illustrated in Figure 2. If a chi-square test is done on each one of these filters (against their respective shadow filters) individual sensor failure isolation is possible. In Table 5 below, an "X" under a filter indicates that the filter has failed the test, a "0" indicates the test passed, and "-" indicates a don't care. A sensor is isolated only in the scenarios listed. Other cases are ambiguous and no action is taken.

Table 5: Chi-Square Failure Isolation

Results of chi-square tests:					Isolate							
F_1^1	F_1^2	F_1^3	F_1^4	F_1^0	F_0^0	F_2^0	F_2^4	F_2^3	F_2^2	F_2^1	Failure	
-	X	X	X	X	X	0	0	0	0	0	0	INS1
0	0	0	0	0	X	X	X	X	X	-	-	INS2
0	X	X	X	X	0	X	X	X	X	0	-	S1
X	0	X	X	X	0	X	X	X	0	X	-	S2
X	X	0	X	X	0	X	X	0	X	X	-	S3
X	X	X	0	X	0	X	0	X	X	X	-	S4

0 - Filter passes chi-square test

X - Filter fails chi-square test

- Irrelevant

The essence of Table 5 is that a sensor failure can be isolated by the chi-square test *only if* all filters that use that sensor fail and all filters that do not use that sensor pass. The only exception to this statement might be when isolating an INS failure (the first two cases in Table 5). It is probably better *not* to wait for filters F_1^1 and F_2^1 to fail. These filters are deprived of

the most accurate aiding sensor, S1, and so they may not be sufficiently sensitive to promptly detect a soft INS failure.

5.4. False Alarms

Since a failure detection results in reconfiguration and long term loss of a sensor, the probability of false alarm must be kept very low; much lower than would be acceptable in a one-time measurement rejection test. Fortunately the voting schemes described in the previous two sections makes this possible without using very high trip levels on the individual tests. For example the INS failure detection by the chi-square test as shown in Table 5 requires the confirmation of five failed chi-square tests and five passed. If these ten individual tests were all independent, with individual false alarm probabilities of α , then the combined probability of a false alarm in either one of the INS's would be:

$$\Pr(\text{fa}) = 2\alpha^5(1-\alpha)^5 \quad (14)$$

Thus a 95% chi-square threshold ($\alpha = .05$) on the individual tests would yield a very small combined false alarm probability of only 5×10^{-7} . The individual thresholds could be increased to 90% or even 75% and still have a combined false alarm probability of .00001 or .0005 respectively.

A case of special interest is the submarine scenario, where only one aiding sensor (the Speed Log) may be available. In this case only two failed and one passed chi-square test are available to detect an INS failure. Thus individual test levels of 95% and 90% yield combined false alarm probabilities of .0048 and .018 respectively.

6. SIMULATION RESULTS: CHI-SQUARE FDI

This section presents some early simulation results performed to investigate the effectiveness of these chi-square tests. Different levels of accelerometer and gyro failures were simulated at different points in time, and different subsets of the state vector were used in the chi-square test. These were evidently among the factors determining effectiveness.

As would be expected, the detectability of a failure depends strongly on the magnitude of the failure simulated, since this determines the size of x_m and hence of $\delta\dot{x}$ and k (via Eqs. 8 and 12). For the purpose of this report, moderately accurate INS's are simulated (1 nmi/hour) and the inertial sensor failures are about eight times larger than specified as normal.

The temporal sensitivity arises from two independent factors:

- the presence of maneuvers, and
- the time since initialization of the shadow filter.

The first factor influences the observability of the inertial sensor errors, and hence the degree to which the corresponding states grow. The second factor is probably a result of the growth in the covariance of the shadow filter state estimates, particularly for the position states.

Two different subsets of the state vector were tested; one with seven elements and one with five. As will be shown below, the difference in effectiveness was not very significant.

Figures 3 to 8 illustrate these effects by showing six different cases. The same 5.5 hour trajectory was used in each case. Each figure shows the test statistic k^2 (from Eq. 12) for the two top level navigation filters, each of which is using the same aiding sensors but a different INS. For each figure one INS has a sensor failure (either accelerometer or gyro) while the other INS has no failure. In all cases the failure was detected, however it is the speed and decisiveness of detection that is of interest.

In Figures 3 to 6 the failed sensor was the x-axis accelerometer. In Figures 7 and 8 the failed sensor is the x-axis gyro. The simulated failures were of the "ramp to constant" type, whereby an additional unmodeled error (accelerometer bias or gyro drift rate) is introduced at a zero level at the "failure time" and increases linearly over an interval (500 seconds) to a maximum value, thereafter remaining fixed. The maximum values used for these failures were $1000 \mu g$ for the accelerometer and 0.1 deg/hr for the gyro.

Figures 3 and 4 can be considered the baseline results. These both illustrate the prompt detection of an accelerometer failure which occurs immediately before a ship maneuver, using a seven degree of freedom (dof) chi-square test. For Figure 3 only Speed Log aiding was used, while for Figure 4 both GPS and Speed Log aiding was used. From this we can see that the Speed Log alone was quite adequate.

Figure 5 illustrates the same situation as in Figure 3 (accelerometer failure before a maneuver, Speed Log aiding) except that a five dof chi-square test was used. This seems to be a slightly less decisive test in this case.

Figure 6 is also the same situation as Figure 3 (accelerometer failure, Speed Log aiding, seven dof) except that the failure takes place 50 minutes before the maneuver. This significantly delays the failure detection.

Figures 7 and 8 illustrate gyro failure detection using

the seven dof chi-square test.

In all of these figures, 3 to 8, the difference between the test statistic k^2 from the failed and unfailed filters is quite dramatic. As discussed in 5.4 lower thresholds could probably be used to improve response time.

7. CONCLUSIONS

The techniques outlined in this paper (multi-filter residual testing and chi-square testing) show considerable promise in being capable of promptly and reliably detecting subtle inertial sensor faults, even with Speed Log aiding only. However further investigation is required to determine the best choice of filters and the most effective FDIR techniques. In particular, the choice of state vector elements for use in the chi-square testing, the choice of confidence levels for use in the individual tests, the choice of individual tests to combine for final failure detection, and so on. Longer simulations are also needed to determine whether or not the method continues to lose sensitivity as the shadow filter covariance grows.

8. REFERENCES

- Bird, J.S., McMillan, J.C. and Arden, D.A.G. (1992). "A Highly Fault Tolerant, Dual Inertial Integrated Navigation System (DIINS)," Proceedings of the National Technical Meeting of the Institute of Navigation, pp 235-244.
- Brumback, B.D., and M.D. Srinath (1987). "A Fault-Tolerant Multisensor Navigation System Design," *IEEE Trans. Aerospace and Electronic Systems, AES-23*, No. 6, pp. 738-756.
- Carlson, N.A. (1987). "Federated Square Root Filter for Decentralized Parallel Processes," Proc. *National Aerospace and Electronics Conf.*, pp. 1448-1456. Also appears in *IEEE Trans. Aerospace and Electronic Systems, AES-26*, No. 3, 1990, pp. 517-525.
- Kerr, T. (1987). "Decentralized Filtering and Redundancy Management for Multisensor Navigation," *IEEE Trans. Aerospace and Electronic Systems, AES-23*, No. 1, pp. 83-119.
- Magill, D.T. (1965). "Optimal Adaptive Estimation of Sampled Stochastic Processes," *IEEE Trans. Automatic Control, AC-10*, No. 4, pp. 434-439.
- McMillan, J.C. (1990). "MINS-B II: A Marine Integrated Navigation System," in *Navigation Land, Sea Air & Space*, M. Kayton editor, IEEE Press, New York, pp. 161-171.
- Widnall, W. (1987) private communication.

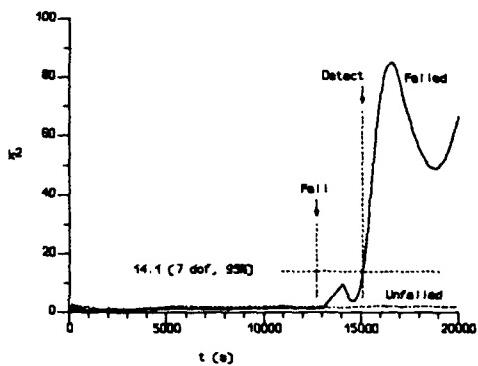


Figure 3. Accelerometer Failure, Speed Log Aiding,
7 dof Chi-Square Test

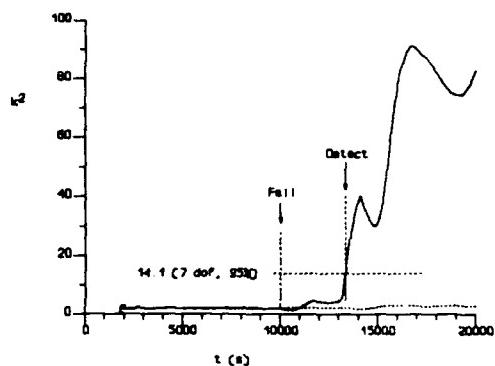


Figure 6. Accelerometer Failure, Speed Log Aiding,
7 dof Chi-Square Test, Delayed Maneuver

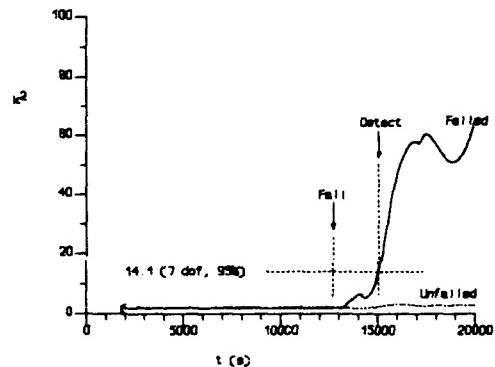


Figure 4. Accelerometer Failure, GPS & Log Aiding,
7 dof Chi-Square Test

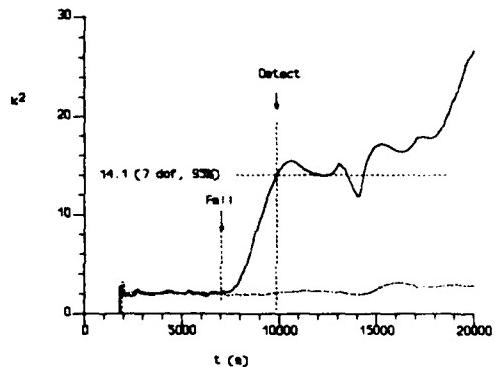


Figure 7. Gyro Failure, Speed Log Aiding, 7 dof Chi-Square Test

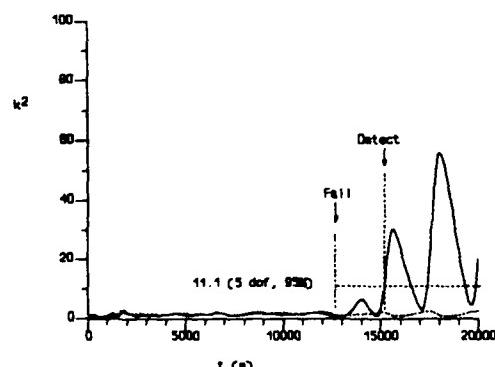


Figure 5. Accelerometer Failure, Speed Log Aiding,
5 dof Chi-Square Test

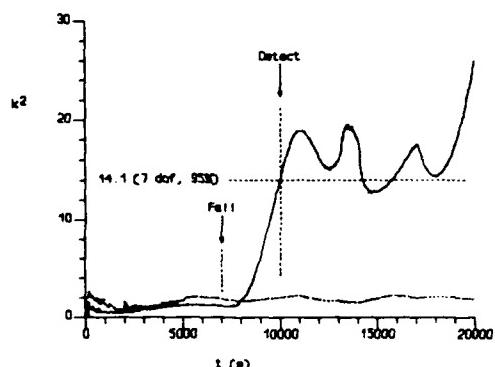


Figure 8. Gyro Failure, GPS & Log Aiding, 7 dof
Chi-Square Test

Toward Achieving Global Sole Means Radionavigation Systems

F. van Graas
 Avionics Engineering Center
 Department of Electrical and Computer Engineering
 361 Stocker Center
 Ohio University
 Athens, Ohio 45701, U.S.A.

SUMMARY

This paper briefly reviews tentative requirements for global, earth-referenced sole means of navigation systems with emphasis on integrity and availability. These requirements can be allocated to integrated navigation system architectures based on for instance GPS, GLONASS, VOR/DME, TACAN, Omega, Chayka, and Loran-C. Fault detection and isolation techniques (FDI) for integrated radionavigation systems are presented. The FDI algorithm provides a protection radius with a specified confidence level as a function of measurement geometry and algorithm requirements. This is followed by a case study of integrated GPS/Loran-C.

LIST OF SYMBOLS AND ACRONYMS

b	Baseline between two stations	Q	Orthonormal matrix
\underline{b}	Measurement bias vector	Q_p	Matrix which columns span parity space
b_i	Minimum detectable bias in measurement space	Q_i	Matrix containing the first m rows of Q^T
c	Speed of light	r_i	Norm of the projection of the parity vector on the i th measurement axis
COV	Covariance	R	Upper triangular matrix
DME	Distance Measuring Equipment	R_i	Range to station i
DOP	Dilution of Precision	R_p	Protection Radius
\underline{e}	Measurement noise vector	RNP	Required Navigation Performance
$\underline{\epsilon}$	Least squares residual vector	σ	Measurement noise standard deviation
erfc	Complementary error function	s_i	Distance between the parity vector and the i th measurement axis
$E\{\cdot\}$	Expected value operator	θ	Bearing
FANS	Future Air Navigation Systems	T_d	Detection threshold in parity space
FDI	Fault detection and isolation	TD	Time difference
GNSS	Global Navigation Satellite System	TACAN	Tactical Air Navigation
GLONASS	Global Navigation Satellite System	U	Matrix containing the first m rows of R
GPS	Global Positioning System	VAR	Variance
H	Data matrix	VOR	Very-High Frequency Omnidirectional Range
HDOP	Horizontal Dilution of Precision	W	Positive definite weighting matrix
I	Identity matrix	\underline{x}	User state vector
ICAO	International Civil Aviation Organization	Δx	Horizontal position error vector
Loran	Long Range Navigation	XDOP	X direction DOP
μ_i	Bias in parity scalar caused by measurement i	y	Measurement vector
μ_m	Minimum required bias in parity space for detection	YDOP	Y-direction DOP
m	Number of elements in \underline{x}		
\underline{m}	Vector containing the i th column of Q_p		
n	Number of elements in \underline{y}		
p	Parity vector		
P_A	Probability of an alarm		
P_D	Probability of detection		
P_{MD}	Probability of a missed detection		
q	Transpose of the first row of Q_p		

1. SOLE MEANS OF NAVIGATION

The operational definition of a sole means air navigation system is provided by the 1990 Federal Radionavigation Plan (Ref 1):

"An approved navigation system that can be used for specific phases of air navigation in controlled airspace without the need for any other navigation system."

This general definition can be translated into specific requirements for a set of parameters, which includes system accuracy, reliability, availability, integrity, coverage, etc. Recently, it has become apparent that the interpretation of these parameters is not clear, especially if these parameters are to represent a global sole means system. For example, GPS coverage is a function of both location and time. The impact of a single-satellite failure cannot be readily compared with the failure of, for instance, a VOR station. This matter is further complicated by the need to integrate different navigation systems in order to satisfy candidate availability and continuity of service requirements. These requirements are not anticipated to be met by for instance

a single satellite navigation system. To accommodate the idea of a navigation function rather than emphasizing a single system, ICAO FANS' GNSS Sub-Group recently defined sole means as follows (Ref 2):

"A means of navigation of the aircraft where position determination is provided by a system which satisfies the required navigation performance (RNP) for a particular phase of flight."

The required navigation performance in this definition is given by:

"RNP is a measure of the navigation system performance within a defined airspace including the operating parameters of the navigation systems used within that airspace."

RNP can be defined using four main parameters: accuracy, availability, integrity, and continuity of service. At this time, only tentative values are available for each of these parameters. For example, the current horizontal positioning accuracy requirements for the nonprecision approach is on the order of 556 meters (95%), while the GNSS Sub-Group recently proposed 150 meters (Ref 2). For the purpose of this paper, it is only necessary to have a general understanding of the requirements. The design of the integrated navigation solution is such that the actual requirement is an input to the solution algorithm.

Once the accuracy requirement is satisfied, two risks exist:

- 1.) Continuity of navigation function risk, which is directly related to the probability of an outage;
- 2.) Navigation integrity risk, which results from an undetected navigation system error.

For a global sole means navigation system, both risks should be on the order of a small multiple of 10^{-9} per flight hour. This results in an unavailability requirement of less than one minute per year, and a probability of loss of integrity of approximately 10^{-9} per flight hour in the absence of radar surveillance (Refs. 2, 3). The probability of loss of integrity could be increased if radar surveillance is available. It is important to realize that these type of requirements must be satisfied through design rather than through simulation. For example, if one simulation would take one second, then it would require more than 3000 years of computer time to perform 10^{11} simulations to verify with some confidence the 10^{-9} probability.

For radionavigation systems, loss of integrity occurs when the horizontal radial position error exceeds a specified alarm threshold for the phase of flight in progress without the integrity alarm being annunciated. Two loss of integrity scenarios must be considered; scenario (1): the horizontal radial position error is exceeded, but no signal malfunction occurred and no fault was detected; and scenario (2): the horizontal radial position error is exceeded because of a signal malfunction, but no fault was detected.

Consider scenario (1), since no signal malfunction is

present, the detection algorithm cannot distinguish between "normal" measurement errors and measurement errors that would cause the horizontal radial position error to exceed the alarm threshold. All one can do to satisfy the integrity requirement at all time and space points is to define a protection radius R_p such that statistically, the probability of exceeding R_p is less than approximately 10^{-9} . Thus, R_p is solely a function of the position fixing statistics of the navigation systems.

Scenario (2) has a much smaller probability of occurrence than scenario (1). If the signal malfunction would cause the horizontal position error to exceed the alarm threshold, then, the detection algorithm should detect this malfunction with a detection probability of $P_D = 1 - (10^9 / P_m)$, where P_m is the probability of a signal malfunction per flight hour. Equivalently, the probability of a missed detection, P_{MD} , should be less than $(1-P_D)$. Typical values for P_{MD} lie in the range of 10^{-3} to 10^{-5} . Note that the detection probability is on a per sample basis, where each sample consists of a set of measurements taken at a specific time.

Currently, integrity for a global sole means system must be accomplished by the airborne system, with little help from external sources. Therefore, it is imperative that a complete understanding of receiver autonomous integrity techniques exists, such that the availability and continuity of service performances can be evaluated. The next section shows that measurements from different radionavigation systems can be integrated into a generic navigation solution, which will then be used to develop the FDI algorithm.

2. INTEGRATED NAVIGATION SOLUTION

Radionavigation system measurements can be expressed in terms of range and bearing observations:

$$\text{Range: } R_i = \sqrt{(X - X_i)^2 + (Y - Y_i)^2 + (Z - Z_i)^2} \quad (1)$$

$$\text{Bearing: } \theta_i = \tan^{-1} \left(\frac{X - X_i}{Y - Y_i} \right)$$

where (X, Y, Z) is the three-dimensional user position and (X_i, Y_i, Z_i) is the position of transmitting station i . R_i is the geometric range between the user and station i , and the bearing is the angle between the user and station i with respect to North. Using the basic range and bearing observations, measurements from other navigation systems can be expressed as well. For instance, a hyperbolic line-of-position is obtained by measuring the time difference (TD) between the times of arrival of signals from two different transmitting stations

$$\text{TD: } TD_{ij} = \frac{b - R_i + R_j}{c} \quad (2)$$

where TD_{ij} is the time difference observation for stations i and j ; b is the geometric distance between the two stations; c is the speed of propagation of the radiowaves; and R_i & R_j are given by equation (1). Note that for terrestrial systems such as Loran-C and Omega, the signals travel great-circle paths. To compensate for this, the transmitter locations are projected onto a locally-level plane with respect to the user estimate at distances equal to the great-

circle distance to the transmitters.

Next, the measurement equations are linearized to arrive at a generic navigation solution. An *a priori* estimate of the user position is used to form a Taylor series expansion, of which only the first order terms are kept.

$$\begin{aligned} R_i &= \hat{R}_i + \frac{\partial R_i}{\partial X}|_{(\hat{x}, \hat{y}, \hat{z})} \delta X + \\ &+ \frac{\partial R_i}{\partial Y}|_{(\hat{x}, \hat{y}, \hat{z})} \delta Y + \frac{\partial R_i}{\partial Z}|_{(\hat{x}, \hat{y}, \hat{z})} \delta Z \end{aligned} \quad (3)$$

The *a priori* position estimate is used to calculate the estimate of the distance to the station (\hat{R}_i). Equation (1) can now be linearized as follows

$$\delta R_i = \left(\frac{\hat{X} - X_i}{\hat{R}_i} \frac{\hat{Y} - Y_i}{\hat{R}_i} \frac{\hat{Z} - Z_i}{\hat{R}_i} \right) \begin{pmatrix} \delta X \\ \delta Y \\ \delta Z \end{pmatrix} \quad (4)$$

A similar procedure is used to linearize the bearing and time difference equations:

$$\delta \theta_i = \left(\frac{\hat{Y} - Y_i}{\hat{R}_i^2} \frac{X_i - \hat{X}}{\hat{R}_i^2} 0 \right) \begin{pmatrix} \delta X \\ \delta Y \\ \delta Z \end{pmatrix} \quad (5)$$

$$\delta TD_i = \frac{1}{c} \left(\frac{\hat{X} - X_i - \hat{X} - X_i}{\hat{R}_i} \frac{\hat{Y} - Y_i - \hat{Y} - Y_i}{\hat{R}_i} \frac{\hat{Z} - Z_i - \hat{Z} - Z_i}{\hat{R}_i} \right)^T \begin{pmatrix} \delta X \\ \delta Y \\ \delta Z \end{pmatrix} \quad (6)$$

Equations (4) through (6) relate a change in the user position to changes in range, bearing, or time difference measurements. In general, equations (4) - (6) can be written as:

$$\delta y_i = h_i \begin{pmatrix} \delta X \\ \delta Y \\ \delta Z \end{pmatrix} \quad (7)$$

where y_i is a measurement, and h_i is a row vector corresponding to that measurement. If all the measurements are included, equation (7) becomes

$$\delta y = H \delta x \quad (8)$$

where y is a vector containing the measurements and x is the user state vector. H is a matrix containing data related to the geometry of the transmitting stations with respect to the user, as given by the row vectors h_i . Equation (8) can be used to iteratively solve for the user state vector.

If an unknown clock offset exists in for instance the range measurement, then another state is included in x . The measurement is then called a pseudorange:

$$PR_i = R_i + cB \quad (9)$$

where B is the unknown clock offset, and c is the speed of propagation of the signal. Linearizing equation (9) results in a slightly different measurement equation:

$$\delta PR_i = \left(\frac{\hat{X} - X_i}{\hat{R}_i} \frac{\hat{Y} - Y_i}{\hat{R}_i} \frac{\hat{Z} - Z_i}{\hat{R}_i} 1 \right) \begin{pmatrix} \delta X \\ \delta Y \\ \delta Z \\ c \delta B \end{pmatrix} \quad (10)$$

More state variables can be added as necessary to solve for other unknowns, such as velocity or acceleration. Hybrid GPS/Loran-Cor GPS/GLONASS, for instance, require that two clock offsets be solved for. Thus, five measurements are required to solve for the user position.

3. FAULT DETECTION AND ISOLATION (FDI)

The concept of FDI is identical to well-known fault detection and isolation techniques for redundant inertial navigation systems and multisensor navigation systems (Refs. 4-8). Recently, these techniques have been applied to GPS by several authors (see for instance Refs. 9, 10). For sole means navigation, two estimators should be used in parallel to achieve FDI:

1. A recursive estimator, typically a Kalman filter, which uses the history of the measurement data to assess the reasonableness of new measurement data. This estimator does not require redundant measurements.
2. A least squares batch estimator which does not rely on the measurement history, but it requires at least one redundant measurement.

The recursive estimator is used to detect and isolate rapidly growing measurement errors by inspecting the estimator residuals, which are the differences between the actual measurements and the predicted measurements based on the history of the measurements. If for instance a measurement residual would be outside the residual interval $[-6.1\sigma, 6.1\sigma]$, where σ is the standard deviation of normally distributed measurement noise, then the probability of this event would be less than 10^{-9} for a static user. This event would be extremely unlikely and the integrity alarm would be raised. In the presence of user dynamics, the residual interval would be increased consistent with the user dynamics. Note that isolation of the faulty measurement is achieved as well.

A least squares batch estimator is used to detect slowly growing measurement errors which go undetected by the recursive estimator. The use of the least squares estimator for this purpose is justified, because:

Independent of the type of estimator used, the position bias error caused by measurement bias errors will always converge to the position bias

error of the least squares estimator (Ref 11). Generally, the time constant of convergence is short compared to the slow error growth of a difficult to detect measurement error.

All information about the inconsistencies in the measurement data is contained in the residuals of the least squares estimator. These residuals are the differences between the actual pseudorange measurements and those predicted based on the least squares position solution and the known transmitter coordinates. The least squares residual vector, $\underline{\epsilon}$, is obtained by projecting the pseudorange measurement vector onto a space perpendicular to the estimation space spanned by the columns of the data matrix. Thus, $\underline{\epsilon}$ is orthogonal to the estimation space. The space in which $\underline{\epsilon}$ lies is also referred to as the parity space. At this point, it is imperative to realize that the least squares residual vector used for FDI lies in parity space and not in the horizontal positioning plane. Therefore, to have a meaningful fault detection statistic, the residual vector must be related to the horizontal radial position error. This is discussed in detail in the next section.

4. GENERIC FDI ALGORITHM

This section presents a least squares fault detection algorithm which extends the concepts presented in references 6, 7, and 12, which were also employed in reference 13, independently.

The linearized relation between changes in the measurements and the corresponding change in the user state vector is given by

$$\underline{\delta y} = H \underline{\delta x} \quad (11)$$

where $\underline{\delta y}$ is a n-by-1 vector containing the changes in the measurements to n sources, $\underline{\delta x}$ is the change in the m-by-1 user state vector, and H is a n-by-m data matrix, see equation (8).

The data matrix H can be decomposed into the product of a real orthonormal matrix Q and an upper triangular matrix R using a "QR" factorization (Ref 14),

$$H = Q R \quad (12)$$

Substituting (12) into (11) and pre-multiplying both sides by Q^T yields ($Q^T Q = I$)

$$R \underline{\delta x} = Q^T \underline{\delta y} \quad (13)$$

The rank of R is equal to the rank of H ; therefore, the lower n-m rows of R consist of zeros only. Equation (13) can be divided into two equations

$$U \underline{\delta x} = Q_1 \underline{\delta y} \quad \text{or} \quad \underline{\delta x} = U^{-1} Q_1 \underline{\delta y} \quad (14)$$

$$Q = Q_p \underline{\delta y} \quad (15)$$

where U consists of the first m rows of R , Q_1 consists of the first m rows of Q^T , and Q_p consists of the last n-m rows of Q^T .

Equation (14) relates the change in the measurements to the change in the user state vector, forming the least squares solution. The rows of Q_p and $\underline{\delta y}$ are orthogonal; therefore, the columns of Q_p span the parity space of H .

If the measurements are corrupted by errors, then $\underline{\delta y}$ is replaced by $\underline{\delta y} + \underline{\epsilon} + \underline{b}$, where $\underline{\epsilon}$ is a n-by-1 vector representing zero-mean, normally distributed measurement noise, and \underline{b} is a n-by-1 vector containing bias errors. Normally, $\underline{\epsilon}$ and \underline{b} are unknown, but their components in parity space are known from equation (15):

$$p = Q_p \underline{\epsilon} + Q_p \underline{b} \quad (16)$$

The expected value of the parity vector, p , is

$$E\{p\} = Q_p \underline{b} \quad (17)$$

The covariance matrix of p is

$$\text{COV}\{p\} = E\{p p^T\} = Q_p \text{COV}\{\epsilon\} Q_p^T \quad (18)$$

If the measurement noise is uncorrelated and normally distributed with equal variances, then the covariance matrix of the measurement noise is

$$\text{COV}\{\epsilon\} = \sigma^2 I \quad (19)$$

where I is a n-by-n identity matrix. It then follows that the covariance matrix of p is

$$\text{COV}\{p\} = \sigma^2 I \quad (20)$$

since $Q_p Q_p^T = I$ (the rows of Q_p are orthonormal vectors).

In the absence of bias errors, the parity vector p is a function of measurement noise only. In the presence of bias errors, p also depends on the bias errors.

4.1 One Redundant Measurement

Assume that only one redundant measurement is available, or $n = m + 1$. In this case, the parity space is one-dimensional and Q_p is reduced to a row vector (the vector q is used to denote the transpose of the first row of Q_p), and the parity vector p is reduced to a parity scalar p . In the absence of bias errors, p has a zero-mean normal probability density function given by

$$f_p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x}{\sigma\sqrt{2}}\right)^2} \quad (21)$$

Assume that the failure detection is based on exceeding a detection threshold T_D , then the probability of an alarm is given by

$$P_A = P(|p| > T_D) = \frac{1}{\sigma\sqrt{2\pi}} \int_{T_D}^{\infty} e^{-\left(\frac{x}{\sigma\sqrt{2}}\right)^2} dx \quad (22)$$

which can also be written as

$$P_A = \operatorname{erfc} \left(\frac{T_D}{\sigma\sqrt{2}} \right) \quad (23)$$

where erfc is the complementary error function

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-x^2} dx \quad (24)$$

In the presence of a bias error in measurement i , the absolute value of the parity scalar has a normal distribution with a mean value of

$$\mu_i = |\mathbf{q} \cdot \mathbf{b}| \quad (25)$$

Only the i th elements of \mathbf{q} and \mathbf{b} will contribute to the mean. The probability density function of p is then given by

$$f_p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu_i}{\sigma}\right)^2} \quad (26)$$

Given the detection threshold T_D , the probability of a missed detection is given by

$$P_{MD} = P(|p| \leq T_D) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-T_D}^{T_D} e^{-\left(\frac{x-\mu_i}{\sigma}\right)^2} dx \quad (27)$$

The contribution of the integral between $-\infty$ and $-T_D$ can be neglected, such that equation (27) can be approximated by

$$P_{MD} = \frac{1}{2} \operatorname{erfc} \left(\frac{\mu_i - T_D}{\sigma\sqrt{2}} \right) \quad (28)$$

Equations (23) and (28) provide the performance of the fault detection algorithm in terms of probability of an alarm and probability of a missed detection as a function of:

- Detection threshold T_D ;
- Measurement noise standard deviation σ ;
- Expected value μ_i of the absolute value of the parity scalar p resulting from a measurement bias error in measurement i .

Obviously, the detection algorithm cannot detect a measurement bias error smaller than the level of measurement noise, since the detection threshold T_D must be set high enough to satisfy the requirement for the alarm probability, see equation (23).

Given the required probability of an alarm and the measurement noise standard deviation, the detection threshold is obtained from equation (23)

$$T_D = \sigma\sqrt{2} \operatorname{erfc}^{-1}(P_A) \quad (29)$$

Next, given the probability of a missed detection, the measurement noise standard deviation and the detection threshold, the minimum required expected value, μ_M , of

the parity scalar is obtained from equation (28)

$$\mu_M = T_D + \sigma\sqrt{2} \operatorname{erfc}^{-1}(2P_{MD}) \quad (30)$$

Since $\mu_i = |\mathbf{q} \cdot \mathbf{b}|$, equation (30) implies that the probability of a missed detection is only satisfied if the measurement bias error gives rise to a bias in the absolute value of the parity scalar greater than or equal to μ_M . The vector \mathbf{q} is known from the measurement geometry; therefore, for each measurement i ($i = 1$ through n), the minimum bias error b_i required to satisfy the probability of a missed detection is calculated from equation (25).

$$b_i = \frac{\mu_M}{|\mathbf{q}_i|} \quad (31)$$

In other words, given the probability of an alarm, the probability of a missed detection, and the measurement noise standard deviation, it follows that the minimum detectable measurement bias error is a function of the measurement geometry.

Horizontal Radial Position Error

From the probability of an alarm, the probability of a missed detection, the measurement noise standard deviation, and the minimum required measurement bias error it is possible to discuss the horizontal radial position error.

Consider the user state error vector resulting from both measurement noise and bias errors. From equation (14) the user state error vector is

$$\Delta \mathbf{x} = \mathbf{U}^{-1} \mathbf{Q}_x (\mathbf{e} + \mathbf{b}) \quad (32)$$

The expected value of the user state error vector is

$$E(\Delta \mathbf{x}) = \mathbf{U}^{-1} \mathbf{Q}_x \mathbf{b} \quad (33)$$

and the error covariance matrix of $\Delta \mathbf{x}$ is

$$\operatorname{COV}(\Delta \mathbf{x}) = \sigma^2 (\mathbf{U}^T \mathbf{U})^{-1} \quad (34)$$

As shown by equations (32) through (34), the effects of the noise and bias errors can be examined separately.

Assume that the matrices \mathbf{U} and \mathbf{Q}_x are expressed in a locally-level reference frame, then the horizontal components of the expected value and the variance of the user state error vector are given by

$$E(\Delta x_H) = \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \quad (35)$$

$$\operatorname{VAR}(\Delta x_H) = \begin{pmatrix} \sigma_x^2 & X \operatorname{DOP}^2 \\ \sigma_y^2 & Y \operatorname{DOP}^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 \\ \sigma_y^2 \end{pmatrix} \quad (36)$$

where \hat{x}, \hat{y} are the first two components of $E\{\Delta x\}$; XDOP² and YDOP² are the first and second diagonal elements of $(U^T U)^{-1}$, respectively.

First consider the horizontal position bias error resulting from measurement bias errors. Each satellite has a minimum bias error necessary for detection with probabilities P_A and P_{MD} as given by equation (31). Each of these measurement bias errors can be converted into a horizontal position error using equation (33). Next, the worst case measurement error is the b_i which maximizes the norm of the horizontal radial position error

$$R_{bias} = \max (\sqrt{\hat{x}_i^2 + \hat{y}_i^2}) \quad (37)$$

The detection algorithm now guarantees that a measurement bias will be detected with the required probabilities P_A and P_{MD} if it contributes to a horizontal radial position error of R_{bias} or greater.

Next, consider the horizontal position error resulting from measurement noise only. The components x and y of the horizontal position error have a bivariate normal probability density function (Refs. 15, 16). Although possible using numerical integration techniques, the calculation of the radius of a circle which contains a certain percentage of the position solutions is rather complicated.

To develop a minimum algorithm, only an upper bound for the protected area is determined as follows. Assume that x and y are independent, then a circle with radius R_{noise} protects the estimated horizontal position with a probability P_T of exceeding R_{noise} .

$$R_{noise} = \sigma\sqrt{2} \operatorname{erfc}^{-1}(P_T) \text{HDOP} \quad (38)$$

where HDOP² = XDOP² + YDOP², and σ is the standard deviation of the measurement noise. For example, if $P_T = 10^{-3}$, then $R_{noise} = 6.1 \sigma \text{ HDOP}$.

To calculate the protection radius, two scenarios are considered. Scenario (1): in the absence of measurement bias errors, the protection radius is approximated by equation (38), using a value on the order of 10^3 for P_T . Scenario (2): in the presence of both measurement noise and a measurement bias error, the vector sum of the bias protection radius and the noise protection radius is used. Thus, the protection radius is approximated by

$$R_p = R_{bias} + \sigma\sqrt{2} \operatorname{erfc}^{-1}(P_{MD}) \text{HDOP} \quad (39)$$

where R_{bias} is calculated from the maximum possible horizontal radial position error due to a measurement bias error which cannot be detected with the required detection probability. P_{MD} is the probability of a missed detection for the fault detection algorithm. Note that the probability of loss of integrity is the product of the probability that a fault exists and the probability of a missed detection given that a fault exists.

Because it is not known which scenario exists, the fault

detection algorithm must use the largest of the two protection radii given for the two scenarios. In general, the protection radius is driven by equation (39).

4.2 Two or More Redundant Measurements

If more than one redundant measurement is available, the parity vector will contain more than one element. Both fault detection and isolation is now possible. Each element of the parity vector has a normal probability density function as given by equation (21). To maintain the Gaussian statistics, each element should be examined separately. Recall that a bias error in measurement i has components along the axes in parity space, given by column i of Q_p . Each column of Q_p defines a measurement axis in parity space: an error in measurement i will lie along the i th measurement axis in parity space. To maximize the visibility of a bias error, the parity vector is projected onto each of the measurement axes. The norm of the projection is:

$$r_i = \frac{|\mathbf{m}_i \cdot \mathbf{p}|}{|\mathbf{m}_i|} \quad (40)$$

where \mathbf{m}_i is the i th column of Q_p . Each r_i given by equation (40) is gaussianly distributed, since the above operation is equivalent to a parity space rotation (Ref 10). Instead of one detection statistic, n detection statistics are obtained. Therefore, the detection threshold is given by

$$T_D = \sigma\sqrt{2} \operatorname{erfc}^{-1}\left(\frac{P_A}{n}\right) \quad (41)$$

The minimum bias error, b_i , required to satisfy the probability of a missed detection is calculated from

$$b_i = \frac{\mu_M}{|\mathbf{m}_i|} \quad (42)$$

Similar to the case with one redundant measurement, the protection radius, R_p , is calculated from equations (37) through (39).

4.3 Fault Isolation

Isolation of the faulty measurement is performed as follows. Calculate the distance of the parity vector to each of the measurement axes:

$$s_i = |\mathbf{p} - (\mathbf{m}_i \cdot \mathbf{p})\mathbf{m}_i| \quad (43)$$

If only one measurement contains a bias, then the smallest s_i corresponds to the faulty measurement (maximum likelihood approach). The probability of a wrong isolation can be calculated by determining the probability that the parity vector is at a distance s_i with respect to a measurement axis. Furthermore, the history of the distances s_i can also be used to determine the faulty measurement. For example, if measurement number one has a ramp error of 5 m/s, then the parity vector will move in a direction parallel to measurement axis one. Examining the time history of s_i allows for early detection and isolation.

4.4 Additional Considerations

So far, all measurements were assumed to have the same statistics. If this is not the case, then the FDI algorithm should be modified to incorporate a weighting matrix. For instance, to accommodate different measurement variances, equation (11) is left multiplied by a weighting matrix W (Ref 18)

$$W\delta y = WH\delta x \quad (44)$$

The "QR" factorization is now performed on WH and equation (13) is replaced by

$$R\delta x = Q^T W\delta y \quad (45)$$

and the parity vector is obtained from

$$p = Q_p W(c + b) \quad (46)$$

In general, W could be derived from the measurement noise covariance matrix, but in most applications it is sufficient to simply use a diagonal matrix, where the diagonal elements are the inverse of the measurement noise standard deviation. If W is correctly selected, then the elements of $W\delta y$ have unit standard deviation.

Exhaustive testing of the integrity algorithm requires a large amount of calculations; e.g. to verify the statistics of the horizontal radial position error, a minimum of 10^{11} simulations are required. Fortunately, because of the Gaussian statistics of the algorithm described above, it is possible to test the algorithm using increased measurement errors. The performance of the FDI algorithm can then be compared with the predicted performance considering the effects of the increased measurement errors (see also Ref 17).

5. GPS/LORAN-C CASE STUDY

5.1 Prototype GPS/Loran-C Receiver

The block diagram of the hardware configuration of the prototype hybrid GPS/Loran-C receiver is shown in Figure 1. A four-channel GPS receiver (Motorola, model Eagle)

and an eight-channel Loran-C receiver (Advanced Navigation, Inc., Model 5300), both employing continuous tracking, are used to collect GPS and Loran-C data. Only the raw measurement data from both receivers is used to determine the position solution. The two receivers are interfaced to a microcomputer (model AT) through two serial communication ports. The microcomputer is also interfaced to a course deviation indicator (CDI, model KI 206), through a parallel port, to display the guidance data to the pilot. All of the hardware used is commercially available equipment, except for the interface between the microcomputer parallel port and the CDI instrument, which was designed and implemented at Ohio University. The software modules implemented on the hybrid GPS/Loran-C receiver are executed in real time. The algorithm of the main procedure is given by:

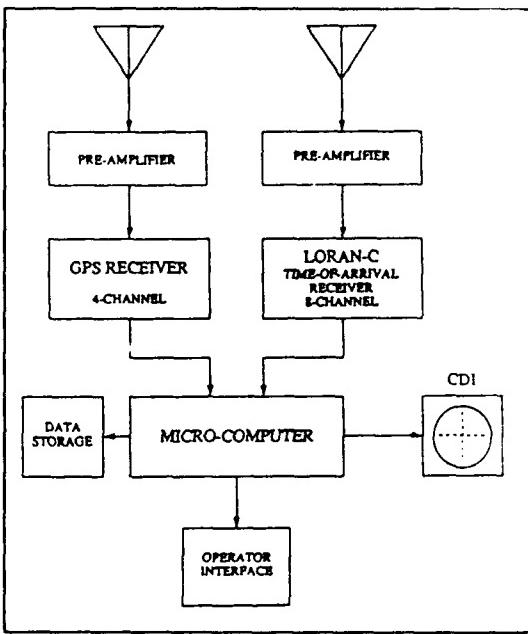


Figure 1. Block diagram of the hardware configuration of the prototype hybrid GPS/Loran-C receiver.

initialization

WHILE in operation

DO once per second

check for keyboard input data
IF keyboard input data
process keyboard data

FIND

check for GPS and Loran-C
measurement data, and request Loran-C
data

IF sufficient data
calculate position and
determine integrity

END

update CDI and status screen
store all relevant data

END

END
system shut-down

During system initialization, the GPS receiver is commanded to send measurement data at a rate of once per second. As soon as GPS data is received, a Loran-C measurement trigger command must be sent to ensure that the Loran-C data is valid at the same time as the GPS data. Following the measurement trigger command, Loran-C data is requested and collected for up to eight receiver channels. All data is verified for validity as indicated by the receivers. If at least five measurements are valid, sufficient data is available for the position calculation. The five measurements are used to solve for three-dimensional position, clock offset with respect to GPS time, and clock offset with respect to Loran-C time. The number of required measurements could be reduced to four if the

GPS and Loran-C receivers measure the time-of-arrivals with respect to the same clock, and if the hardware delays of both receivers are known. A minimum of seven valid measurements are required to execute the FDI algorithm.

The hybrid position solution is based on a least-squares estimator. Since the measurements from GPS and Loran-C are equally weighted, the accuracy of the hybrid system will be mostly determined by the Loran-C measurements. For this effort, standard Loran-C propagation models are used such that the achieved accuracies are representative for current Loran-C receivers. Because of this, the accuracy of the hybrid system will not be as good as that provided by GPS; however, the availability and integrity of the hybrid system exceeds that of GPS by several orders of magnitude. At the same time, the hybrid system accuracies are still well within the current requirements.

5.2 Flight Test Results

Flight test data collected on August 23, 1990 was used for evaluating the FDI algorithm. The flight lasted approximately 52 minutes and was conducted in the vicinity of the Ohio University Airport in Albany, Ohio (Ref 19). A 5 m/s error was simulated in each of the seven measurements during a 500-second portion of the flight.

Although the measurement error reached a value of 2500 meters, the actual horizontal radial position error did not exceed 1000 meters. The trajectories of the parity vectors over time for all seven simulated measurement errors are shown in Figure 2. Each parity trace begins at the same point and extends outward in a direction parallel to the corresponding measurement axis. Note that the axes in Figure 2 are only valid for one particular moment of time. For this example, the measurement axes did not change much during the 500-second run, except for measurement axis two, which rotated rapidly counter-clock-wise. The detection threshold in parity space was set at 950 meters. All measurement errors were correctly detected and isolated, except for measurement 2. However, the position error caused by measurement 2 did not exceed 250 meters, and therefore, the alarm did not need to be raised. If the simulation would have been continued, the error in measurement 2 would also have been detected and isolated correctly. An example of the isolation process is shown in Figure 3, where the distance of the parity vectors with respect to each measurement axis is graphed as a function of time. This particular example involves failing measurement 6. The distance to measurement axis 6 is shown as a solid line in the figure.

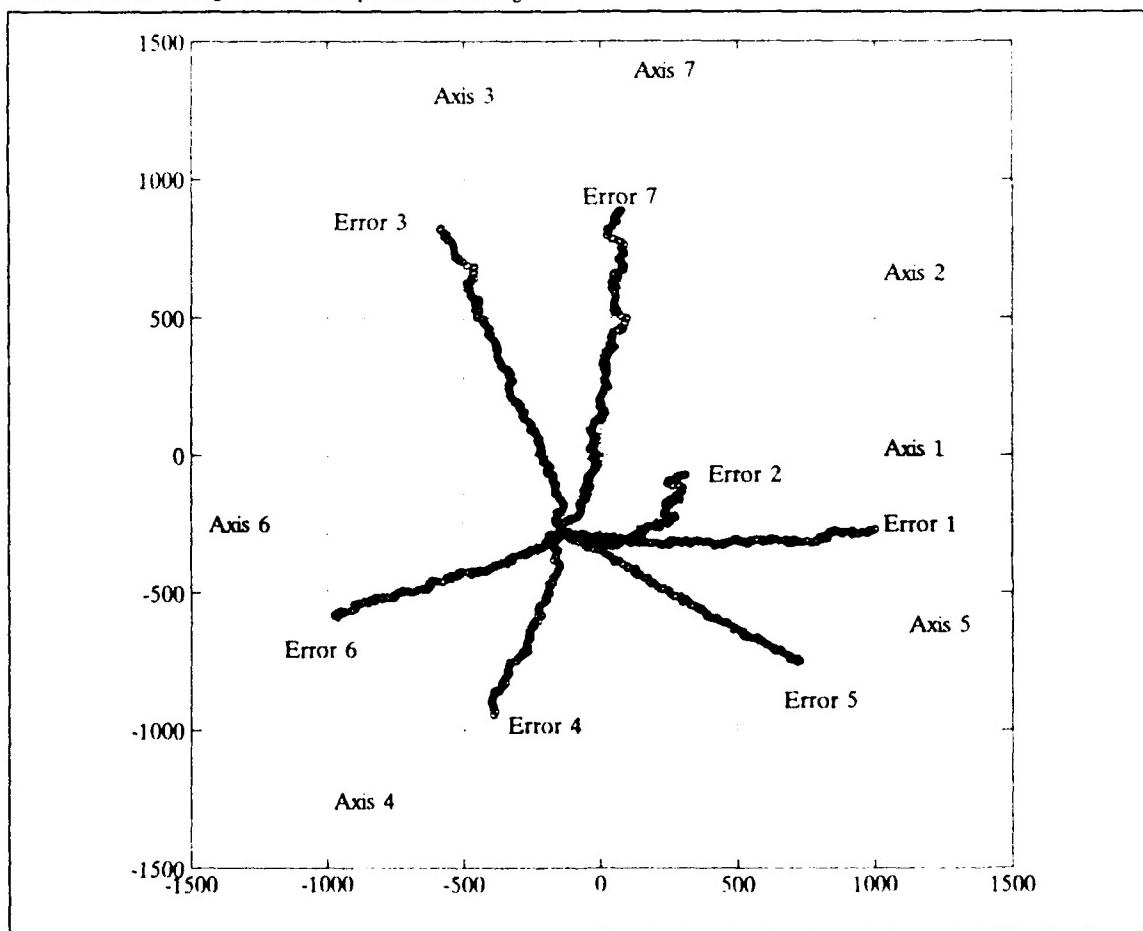


Figure 2 Trajectories of the Party Vectors over Time for a 5 m/s Ramp Error in Each of the Seven Measurements

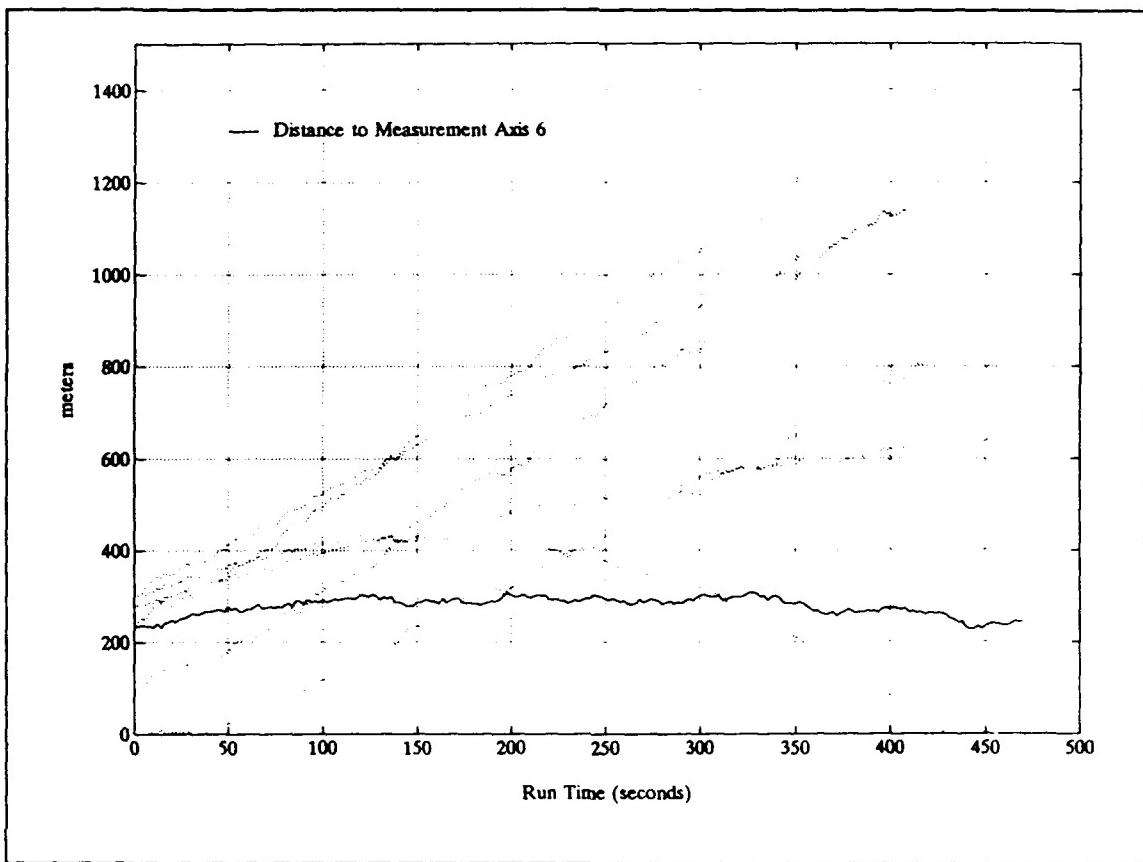


Figure 3. Distances to Each of the Measurements Axes over Time for a 5 m/s Ramp Error in Measurement 6.

Since the distance to measurement axis 6 is fairly constant compared to the distances to the other axes, it follows that measurement 6 is malfunctioning.

6. CONCLUSIONS

Global sole means of navigation systems could be obtained by integrating all available measurement sources and applying FDI techniques. Key to achieving FDI is an algorithm that calculates the protection radius with a pre-specified confidence level as a function of measurement geometry and algorithm requirements. The parity space algorithm outlined in this paper is capable of performing this task.

7. ACKNOWLEDGEMENTS

The fault detection portion of this paper is based on an earlier paper presented at the National Technical Meeting of The Institute of Navigation, Phoenix, Arizona, Jan. 1991. Funding for that work was obtained from the FAA under contract DTRS-57-87-C-00006, TTD-36 and TTD-46. Additional funding for the development of FDI algorithms was obtained from the FAA and NASA Langley Research Center under the Joint University Program for Air Transportation Research, Grant NGR-009-017. Portions of this work have also been published in a Master's Thesis written by P.A. Kline while at Ohio University (Ref 19).

8. REFERENCES

1. U.S. Department of Transportation and U.S. Department of Defense, "1990 Federal Radionavigation Plan", Report No. DOT-VNTSC-RSPA-90-3/DOD-4650.4, Washington, D.C., 1990.
2. ICAO FANS GNSS Sub-Group, Meeting report FANS(II)3-WP/87, Report on Agenda Items 4 and 5, 1992.
3. Anoll, R.K., "Some Thoughts on Reliability and Availability Requirements for GNSS and Other Sole-Means Navigation Systems", RTCA Paper No. 374-92/SC159-367, April, 1992.
4. Wilcox, J.C., "Competitive Evaluation of Failure Detection Algorithms for Strapdown Redundant Inertial Instruments", J. of Spacecraft and Rockets, Vol. 11, July 1974.
5. Gai, E.G., Adams, M.B. and Walker, B.K., "Determination of Failure Thresholds in Hybrid Navigation", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-12, November 1976.

6. Daly, K.C., Gai, E.G. and Harrison, J.V., "Generalized Likelihood Test for FDI in Redundant Sensor Configurations", *J. Guidance and Control*, Vol. 2, Jan.-Feb. 1979.
7. Potter, J.E. and Suman, M.C., "Thresholdless Redundancy Management with Arrays of Skewed Instruments", *Integrity in Electronic Flight Control Systems*, AGARDograph-224, 1977.
8. Kerr, T.H., "Decentralized Filtering and Redundancy Management for Multisensor Navigation", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-23, No. 1, January 1987, pp 83-119.
9. Sturza, M.A., "Navigation System Integrity Monitoring Using Redundant Measurements", *Navigation: Journal of the Institute of Navigation*, Vol. 35, No. 4, Winter 1988-89, pp. 483-501.
10. Brenner, M., "Implementation of a RAIM Monitor in a GPS Receiver and an Integrated GPS/IRS", *Proceedings of the Third International Satellite Division Meeting of the ION*, Colorado Springs, CO, September 1990.
11. Kuhl, M.R., "Ridge Regression Signal Processing Applied to Multisensor Position Fixing", Master's Thesis, Department of Electrical and Computer Engineering, Ohio University, June 1990.
12. Parkinson, B.W., and Axelrad, P., "Receiver Autonomous Integrity Monitoring of the Navigation Service Provided by the Global Positioning System", Report prepared for the U.S. DOT, NVTSC, Contract no. 88-P-82289, May 1989.
13. Brenner, M., "Implementation of a RAIM Monitor in a GPS Receiver and an Integrated GPS/IRS", *Proceedings of the Third International Satellite Division Meeting of the ION*, September 1990.
14. Golub, G.H. and Van Loan, C.F., "Matrix Computations", Second Edition, The Johns Hopkins University Press, Baltimore, MD, 1989.
15. Chin, G.Y., "Two-Dimensional Measures of Accuracy in Navigational Systems", Final Report, no. DOT-TSC-RSPA-87-1, U.S. DOT, NVTSC, March 1987.
16. Torrieri, D.J., "Statistical Theory of Passive Location Systems", *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AEC-20, March 1984.
17. Van Graas, F. and Farrell, J.L., "Receiver Autonomous Integrity Monitoring (RAIM): Techniques, Performance and Potential", *Proceedings of the 47th Annual Meeting of the ION*, Williamsburg, VA, June 10-12, 1991.
18. Lawson, C.L. and Hanson, R.J., "Solving Least Squares Problems", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
19. Kline, P.A., "Fault Detection and Isolation for Integrated Navigation Systems Using the Global Positioning System", Master's Thesis, Department of Electrical and Computer Engineering, Ohio University, Athens, Ohio, November 1991.

REPORT DOCUMENTATION PAGE				
1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document	
	AGARD-CP-525	ISBN 92-835-0693-6	UNCLASSIFIED	
5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 Rue Ancelle, 92200 Neuilly sur Seine, France			
6. Title	INTEGRATED AND MULTI-FUNCTION NAVIGATION			
7. Presented at	the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th—15th May 1992.			
8. Author(s)/Editor(s)	Various		9. Date	
				November 1992
10. Author's/Editor's Address	11. Pages			
				176
12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the back covers of all AGARD publications.			
13. Keywords/Descriptors	Accelerometers Automatic landing Differential GPS Fault tolerance Fiber optic gyro GPS Gyroscopes Hemispherical resonator gyro Integrated GPS/INS Integrated navigation Multifunction navigation Reference systems Ring laser gyro Scene correlation			
14. Abstract	<p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th—15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> — Mission Applications; — Sensors for Integrated Navigation and Multi-Function Reference Systems; — Design of Integrated and Multifunctional Navigation Systems. 			

			AGARD-CP-525
AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages	<p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	<p>AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages</p> <p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	AGARD-CP-525
AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages	<p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	<p>AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages</p> <p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	AGARD-CP-525
AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages	<p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	<p>AGARD Conference Proceedings 525 Advisory Group for Aerospace Research and Development, NATO INTEGRATED AND MULTI-FUNCTION NAVIGATION Published November 1992 176 pages</p> <p>This volume contains the 14 unclassified papers, presented at the Guidance and Control Panel Specialists' Meeting held in Ottawa, Canada from 14th – 15th May 1992.</p> <p>The papers were presented covering the following headings:</p> <ul style="list-style-type: none"> - Mission Applications; - Sensors for Integrated Navigation and Multi-Function Reference Systems; - Design of Integrated and Multifunctional Navigation Systems. <p>ISBN 92-835-0693-6</p>	AGARD-CP-525

NATO OTAN
7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE
Telephone (1) 47 38 57 00 · Telex 610 176
Telex 1147 38 57 99

**DIFFUSION DES PUBLICATIONS
AGARD NON CLASSIFIÉES**

L'AGARD ne détient pas de stocks de ses publications, dans un but de distribution générale à l'adresse ci-dessus. La diffusion initiale des publications de l'AGARD est effectuée auprès des pays membres de cette organisation par l'intermédiaire des Centres Nationaux de Distribution suivants. À l'exception des Etats-Unis, ces centres disposent partout d'exemplaires additionnels, dans les cas contraires, on peut se procurer ces exemplaires sous forme de microtchées ou de photocopies auprès des Agences de Vente dont la liste suit.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum,
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordonnateur AGARD-VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Directeur du Service des Renseignements Scientifiques
Ministère de la Défense Nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Board
Ved Idrætsparken 4
2100 Copenhagen O

ESPAGNE
INTA (AGARD Publications)
Pintor Rosales 34
28008 Madrid

ETATS-UNIS
National Aeronautics and Space Administration
Langley Research Center
M/S 180
Hampton, Virginia 23665

FRANCE
ONERA (Direction)
29, Avenue de la Division Leclerc
92322 Chatillon Cedex

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens 11010

**LE CENTRE NATIONAL DE DISTRIBUTION DES ETATS-UNIS (NASA) NE DETIENT PAS DE STOCKS
DES PUBLICATIONS AGARD ET LES DEMANDES D'EXEMPLAIRES DOIVENT ETRE ADRESSEES DIRECTEMENT
AU SERVICE NATIONAL TECHNIQUE DE L'INFORMATION (NTIS) DONT L'ADRESSE SUIT.**

AGENCES DE VENTE

National Technical Information Service
(NTIS)
5285 Port Royal Road
Springfield, Virginia 22161
Etats-Unis

ESA Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume Uni

Toutes les demandes de microtchées ou de photocopies de documents AGARD (y compris les demandes faites auprès du NTIS) doivent comporter la dénomination AGARD, ainsi que le numéro de série de l'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuillez noter qu'il vaut mieux spécifier AGARD-R pour les AGARD-AR ainsi que de la commande de rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes, ainsi que des résumés des publications AGARD figurent dans les journaux suivants.

Scientific and Technical Aerospace Reports (STAR)
publié par la NASA Scientific and Technical
Information Division
NASA Headquarters (NTH)
Washington DC 20546
Etats-Unis

Government Reports Announcements and Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 22161
Etats-Unis
Accessible également en mode interactif dans l'université de données bibliographiques en ligne du NTIS, et sur CD-ROM.



*L'imprime par Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3JZ*

NATO
OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE

Telephone (1)47.38.57.00 · Telex 610 176
Telefax (1)47.38.57.99

**DISTRIBUTION OF UNCLASSIFIED
AGARD PUBLICATIONS**

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres (except in the United States), but it may not be purchased in Microfiche or Photocopy form from the Sales Agencies listed below.

NATIONAL DISTRIBUTION CENTRES	
BELGIUM	LUXEMBOURG <i>See Belgium</i>
Coordonnateur AGARD — VSL Etat-Major de la Force Aérienne Quartier Reine Elisabeth Rue d'Evere, 1140 Bruxelles	NETHERLANDS Netherlands Delegation to AGARD National Aerospace Laboratory, NLR Kluiverweg 1 2629 HS Delft
CANADA	NORWAY Norwegian Defence Research Establishment Attn. Biblioteket P.O. Box 25 N-2007 Kjeller
Director Scientific Information Services Dept of National Defence Ottawa, Ontario K1A 0K2	PORTUGAL Portuguese National Coordinator to AGARD Gabinete de Estudos e Programas CLAEFA Base de Alfragide Alfragide 2700 Amadora
DENMARK	SPAIN INIA (AGARD Publications) Pintor Rosales 34 28008 Madrid
Danish Defence Research Board Ved Idrætsparken 4 2100 Copenhagen Ø	TURKEY Milli Savunma Başkanlığı (MSB) ARGEL Daire Başkanlığı (ARDL) Ankara
FRANCE	UNITED KINGDOM Defence Research Information Centre Kentigern House 65 Brown Street Glasgow G2 8EX
O.N.E.R.A. (Direction) 29 Avenue de la Division Leclerc 92322 Châtillon Cedex	UNITED STATES National Aeronautics and Space Administration (NASA) Langley Research Center M/S 180 Hampton, Virginia 23665
GERMANY	
Fachinformationszentrum Karlsruhe D-7514 Eggenstein-Leopoldshafen 2	
GREECE	
Hellenic Air Force Air War College Scientific and Technical Library Dekelia Air Force Base Dekelia, Athens TGA 1010	
ICELAND	
Director of Aviation c/o Flugrad Reykjavik	
ITALY	
Aeronautica Militare Ufficio del Delegato Nazionale all'AGARD Aeroporto Pratica di Mare 00140 Pomezia (Roma)	

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW

SALES AGENCIES

National Technical Information Service (NTIS) 5285 Port Royal Road Springfield, Virginia 22161 United States	ESA Information Retrieval Service European Space Agency 10, rue Mario Nikis 75015 Paris France	The British Library Document Supply Centre Boston Spa, Wetherby West Yorkshire LS23 7BO United Kingdom
---	--	--

Requests for microfiches or photocopies of AGARD documents (including requests to NTIS) should include the word 'AGARD' and the AGARD serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Note that AGARD Reports and Advisory Reports should be specified as AGARD-R-nnn and AGARD-AR-nnn, respectively. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR) published by NASA Scientific and Technical Information Division NASA Headquarters (NTT) Washington D.C. 20546 United States	Government Reports Announcements and Index (GRA&I) published by the National Technical Information Service Springfield Virginia 22161 United States (also available online in the NTIS Bibliographic Database or on CD-ROM)
--	---



Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3JZ